STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 141260

TO:	From:
Anish Gupta Location: REM-3C15&3C18 Art Unit: 1654 Wednesday, December 29, 2004 Case Serial Number: 09/823418	Edward Hart Location: Biotech-Chem Library REM-1A55 Phone: 571-272-2512 edward.hart@uspto.gov

Search Notes

Examiner Gupta,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.
Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

141260

STIC-Biotech/ChemLib

From: Gupta, Anish
Sent: Monday, December 27, 2004 2:39 PM
To: STIC-Biotech/ChemLib
Subject: RE: search request

CRF

Serial Number: 09 / 823418 Art Unit: 1654
Room: Remsen 3C15
Mailbox Room: Remsen 3C18

Please search sequence ID No. 2, 3, 4, 5, 6, 7, 8, 13 and 14

Anish Gupta
2-0859
Remsen.

Protein Sequence Searches - 10/8/04

All of the sequence databases on the ABSS have been updated. A change has occurred in the protein databases.

- Two protein databases, SPTREMBL and SwissProt, are now produced as a single, merged database called UniProt.
- Results from UniProt have the file extension **.rup**.
- Sequences in UniProt are identified by the same ID that had been used in SPTREMBL or SwissProt.
- In instances where the database curators have determined that an SPTREMBL record and a SwissProt record represent the same sequence, the two records have been merged into one. Both IDs are present in the record. Any differences found between the two sequences are recorded in the FT (feature table) fields.

If you have any questions regarding these changes or your results, please contact any STIC searcher.

****
STAFF USE ONLY

Searcher:_____
Searcher Phone: 2-
Date Searcher Picked up: 12/29/04
Date Completed: 12/29/04
Searcher Prep/Rev. Time:
Online Time:_____

****
Type of Search
NA Sequence: #_____
AA Sequence :# 9
Structure: #_____
Bibliographic:_____
Litigation:_____
Patent Family:_____
Other:_____

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM: DP
WWW/Internet:_____
Other(Specify):_____

USPTO

STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
571-272-2507 Remsen E01 D86

Voluntary Results Feedback

- *I am an examiner in Workgroup:* ☐ *Example: 1610*
- *Relevant prior art **found**, search results used as follows:*
  - ☐ 102 rejection
  - ☐ 103 rejection
  - ☐ Cited as being of interest.
  - ☐ Helped examiner better understand the invention.
  - ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
  (journal articles, conference proceedings, new product announcements etc.)

- *Relevant prior art **not found**:*
  - ☐ Results verified the lack of relevant prior art (helped determine patentability).
  - ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.

sira
Search and Information Resources Administration

THIS PAGE IS BLANK

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-2
Perfect score: 50
Sequence: 1 TRLTRDRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	50	100.0	10	2	AAY30683	Aay30683 Apo-B100
2	46	92.0	10	2	AAY30682	Aay30682 Apo-B100
3	44	88.0	10	2	AAY30686	Aay30686 Apo-B100
4	44	88.0	10	2	AAY30687	Aay30687 Apo-B100
5	43	86.0	10	2	AAY30685	Aay30685 Apo-B100
6	43	86.0	11	2	AAW57205	Aaw57205 Apo B bin
7	43	86.0	13	2	AAW57207	Aaw57207 Apo B 100
8	43	86.0	15	2	AAW41261	Aaw41261 Apolipopr
9	43	86.0	15	2	AAW96892	Aaw96892 ApoB-100
10	43	86.0	20	6	ABJ37575	Abj37575 Heparin b
11	43	86.0	22	2	AAW57208	Aaw57208 Apo B 100
12	43	86.0	22	2	AAW57209	Aaw57209 Apo B 100
13	43	86.0	34	5	AAE14541	Aae14541 Human apo
14	43	86.0	36	2	AAW96876	Aaw96876 Nucleic a
15	43	86.0	37	2	AAW64587	Aaw64587 Human apo
16	43	86.0	51	2	AAW96845	Aaw96845 Nucleic a
17	43	86.0	343	4	ABB37687	Abb37687 Peptide #
18	43	86.0	343	4	ABG52504	Abg52504 Human liv
19	43	86.0	377	2	AAR72704	Aar72704 Human apo
20	43	86.0	377	2	AAR34031	Aar34031 Sequence
21	43	86.0	2463	8	ADJ57400	Adj57400 Human apo
22	43	86.0	3923	2	AAY31237	Aay31237 Human Apo
23	43	86.0	4536	2	AAW41262	Aaw41262 Apolipopr
24	43	86.0	4536	2	AAW96826	Aaw96826 Amino aci
25	43	86.0	4560	5	AAU98981	Aau98981 Human apo
26	43	86.0	4561	7	ADD48677	Add48677 Human Pro
27	43	86.0	4563	5	AAO15893	Aao15893 Human apo
28	43	86.0	4563	6	ABR40253	Abr40253 Human ali
29	43	86.0	4563	6	ABU79140	Abu79140 Apolipopr
30	43	86.0	4563	7	ADF43408	Adf43408 Apolipopr
31	43	86.0	4563	8	ADH18871	Adh18871 Human apo
32	43	86.0	4563	8	ADH18870	Adh18870 Human apo
33	43	86.0	4563	8	ADO33445	Ado33445 Human apo
34	43	86.0	4563	8	ADO33447	Ado33447 Human apo
35	43	86.0	4590	4	AAU33184	Aau33184 Novel hum
36	42	84.0	10	2	AAY30684	Aay30684 Apo-B100
37	39.5	79.0	11	2	AAY30699	Aay30699 Apo-B100
38	39.5	79.0	11	2	AAY30700	Aay30700 Apo-B100
39	39	78.0	1583	4	AAB59828	Aab59828 Protein #
40	38	76.0	10	2	AAY30690	Aay30690 Apo-B100
41	38	76.0	10	2	AAY30692	Aay30692 Apo-B100
42	38	76.0	10	2	AAY30688	Aay30688 Apo-B100
43	38	76.0	11	2	AAW57206	Aaw57206 Apo B 100
44	38	76.0	11	2	AAW87717	Aaw87717 Analogue
45	38	76.0	11	5	AAE21732	Aae21732 BSMR effe

ALIGNMENTS

RESULT 1
AAY30683
ID AAY30683 standard; peptide; 10 AA.
XX
AC AAY30683;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10
||||||||||
Db 1 TRLTRDRGLK 10

RESULT 2
AAY30682
ID AAY30682 standard; peptide; 10 AA.
XX
AC AAY30682;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 92.0%; Score 46; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.016;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10
|||||:||||
Db 1 TRLTRERGLK 10

RESULT 3
AAY30686
ID AAY30686 standard; peptide; 10 AA.
XX
AC AAY30686;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 88.0%; Score 44; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.04;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10

```
                  ||||| ||||
Db          1 TRLTRSRGLK 10


RESULT 4
AAY30687
ID   AAY30687 standard; peptide; 10 AA.
XX
AC   AAY30687;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             88.0%;  Score 44;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.04;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db          1 TRLTRQRGLK 10


RESULT 5
AAY30685
ID   AAY30685 standard; peptide; 10 AA.
XX
AC   AAY30685;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             86.0%;  Score 43;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.065;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db          1 TRLTRTRGLK 10


RESULT 6
AAW57205
ID   AAW57205 standard; peptide; 11 AA.
XX
AC   AAW57205;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B binding site peptide 2.
XX
```

KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 12; Page 52; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B binding site
CC peptide which can be used as a component of a non-naturally occurring,
CC receptor-competent low density lipoprotein (LDL) particle of the present
CC invention. The LDL particle comprises at least 1 peptide component that
CC has at least 1 binding site for an apo B protein receptor and at least 1
CC lipophilic substituent. Also described in the invention are peptides
CC containing an apo B binding sequence with at least 70% identity with
CC sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC naturally occurring, receptor-competent LDL particles are useful as: (i)
CC drug-targeting vectors for delivering anticancer drugs to cancer cells
CC that express an apo B protein receptor, and (ii) additives for cell
CC culture media especially as growth supplements. Non-naturally occurring,
CC receptor-competent LDL particles do not require the complete apo B
CC sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 11 AA;

Query Match 86.0%; Score 43; DB 2; Length 11;
Best Local Similarity 90.0%; Pred. No. 0.072;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10
||||| ||||
Db 2 TRLTRKRGLK 11

RESULT 7
AAW57207
ID AAW57207 standard; peptide; 13 AA.
XX
AC AAW57207;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide B.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 13 AA;

Query Match 86.0%; Score 43; DB 2; Length 13;
Best Local Similarity 90.0%; Pred. No. 0.086;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10
||||| ||||
Db 3 TRLTRKRGLK 12

RESULT 8
AAW41261
ID AAW41261 standard; peptide; 15 AA.
XX
AC AAW41261;
XX
DT 19-MAY-1998 (first entry)
XX
DE Apolipoprotein B-100 fragment.
XX
KW Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW prothrombinase complex.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9743311-A1.
XX
PD 20-NOV-1997.
XX
PF 09-MAY-1997; 97WO-GB001255.
XX
PR 09-MAY-1996; 96GB-00009702.
XX

PA (UNLO) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI Bruckdorfer KR, Ettelaie C;
XX
DR WPI; 1998-008798/01.
XX
PT Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT used for treating or preventing coagulation, inhibiting angiogenesis,
PT cell differentiation and apoptosis.
XX
PS Disclosure; Page 22; 60pp; English.
XX
CC This sequence is an example of the peptide of the invention. It has the
CC formula (I), or their variants with one or more internal deletions,
CC insertions or substitutions, while retaining anti-coagulant properties of
CC apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC aa. Compositions containing the peptide are used for simultaneous,
CC separate or sequential treatment of cancer, particularly to prevent
CC metastatic spread. They are also used to inhibit thromboplastin-mediated
CC processes, specifically to prevent or reduce blood coagulation (e.g.
CC during or after surgery or in cases of heart attack, stroke etc.) and to
CC inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC which is active as such or as part of a 98-aa peptide, inhibits
CC activation of the prothrombinase complex; and prevents activation of
CC factor VII on the surface of thromboplastin and of platelets by thrombin.
CC It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC smaller than apoB-100, they act more quickly
XX
SQ Sequence 15 AA;

Query Match 86.0%; Score 43; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10
||||| ||||
Db 1 TRLTRKRGLK 10

RESULT 9
AAW96892
ID AAW96892 standard; peptide; 15 AA.
XX
AC AAW96892;
XX
DT 22-APR-1999 (first entry)
XX
DE ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from
CC human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC component of very-low density lipoproteins (VLDL), intermediate density
CC lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC present sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 15 AA;

Query Match 86.0%; Score 43; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRDRGLK 10
||||| ||||
Db 6 TRLTRKRGLK 15

RESULT 10
ABJ37575
ID ABJ37575 standard; peptide; 20 AA.
XX
AC ABJ37575;
XX
DT 10-MAY-2003 (first entry)
XX
DE Heparin binding peptide sequence #28.
XX
KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS Unidentified.
XX
PN WO2003007689-A2.
XX
PD 30-JAN-2003.
XX
PF 22-JUL-2002; 2002WO-US023419.
XX
PR 20-JUL-2001; 2001US-0306726P.
XX
PA (ETHZ-) ETH ZUERICH.
PA (UYZU-) UNIV ZURICH.
XX
PI Hubbell JA, Schoenmakers R, Maynard HD;
XX
DR WPI; 2003-300420/29.
XX
PT Use of a ligand comprising of at least one sulfated or sulfonated amino
PT acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT retinopathy and hypoxia.
XX
PS Disclosure; Fig 2; 79pp; English.
XX
CC The invention relates to a novel ligand for binding a target biomolecule,
CC which comprises a peptide having at least one sulphated or sulphonated
CC amino acid and at least one amino acid chosen from neutral and positively
CC charged amino acids. The novel ligands can be used for the treatment of
CC e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC This sequence represents a heparin binding peptide relating to the

```
CC   invention
XX
SQ   Sequence 20 AA;

  Query Match             86.0%;  Score 43;  DB 6;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 11
AAW57208
ID   AAW57208 standard; peptide; 22 AA.
XX
AC   AAW57208;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide C.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
FT   Modified-site   22
FT                   /note= "attached to cholesterol"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             86.0%;  Score 43;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 12
AAW57209
ID   AAW57209 standard; peptide; 22 AA.
XX
AC   AAW57209;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide D.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             86.0%;  Score 43;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
```

```
Db          7 TRLTRKRGLK 16


RESULT 13
AAE14541
ID   AAE14541 standard; peptide; 34 AA.
XX
AC   AAE14541;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human apoB-100 derived peptide p62.
XX
KW   Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW   cardiovascular disease; coronary heart disease; pre-eclampsia;
KW   non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW.  peptide p62.
XX
OS   Homo sapiens.
XX
PN   WO200206314-A2.
XX
PD   24-JAN-2002.
XX
PF   18-JUL-2001; 2001WO-GB003212.
XX
PR   18-JUL-2000; 2000GB-00017641.
XX
PA   (ARKT-) ARK THERAPEUTICS LTD.
XX
PI   Narvanen O,  Yla-Herttuala S;
XX
DR   WPI; 2002-179777/23.
XX
PT   New peptide useful in enzyme immunoassays for detecting oxidized low
PT   density lipoprotein which is a marker of coronary heart disease and other
PT   cardiovascular diseases, has affinity for oxidized low density
PT   lipoprotein.
XX
PS   Claim 6; Page 5; 21pp; English.
XX
CC   The invention relates to peptides having affinity for oxidised low
CC   density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC   is useful in an immunoassay to determine the presence, and optionally,
CC   the amount of antibodies in a sample, having affinity for oxLDL.
CC   Preferably immobilised peptide is useful for measuring the amount of
CC   autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC   from a patient for evaluating the risk of coronary heart diseases, other
CC   cardiovascular diseases, and several other disorders such as
CC   periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC   endothelial dysfunction. The peptide of the invention is stable, can be
CC   synthesised easily without the need to isolate proteins from a patient's
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             86.0%;  Score 43;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db         25 TRLTRKRGLK 34


RESULT 14
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Fig 12C; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 36 AA;

  Query Match             86.0%;  Score 43;  DB 2;  Length 36;
  Best Local Similarity   90.0%;  Pred. No. 0.26;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db         11 TRLTRKRGLK 20


RESULT 15
AAW64587
ID   AAW64587 standard; peptide; 37 AA.
XX
AC   AAW64587;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Human apolipoprotein peptide fragment #1.
XX
KW   Factor V; human; detection; protein function; blood coagulation; apo;
KW   fat metabolism; Leyden mutation; deep vein thrombosis; apolipoprotein;
KW   Alzheimer's disease; 5,10-methylenetetrahydrofolate reductase; prion;
KW   hypercysteinaemia; factor VII; cardiovascular disease; pathogen; virus.
XX
OS   Homo sapiens.
XX
PN   EP857973-A2.
```

XX
PD 12-AUG-1998.
XX
PF 12-JAN-1998; 98EP-00890007.
XX
PR 13-JAN-1997; 97AT-00000044.
XX
PA (IMMO) IMMUNO AG.
XX
PI Moritz B, Kiessig S, Lang H, Schenk V;
XX
DR WPI; 1998-416142/36.
XX
PT Detecting or quantifying mutant protein in presence of wild-type protein
PT comprises reaction with ligand - used to detect mutant blood coagulation
PT factors or apolipoproteins for diagnosing risk of thrombosis.
XX
PS Example 2; Page 9; 18pp; German.
XX
CC AAW64587 and AAW64588 are fragments of human apolipoprotein which are
CC used with Factor V protein fragments in a novel method to detect the
CC presence of a mutated protein in a sample that may also contain the
CC corresponding wild-type protein. The method is used to detect mutations
CC that alter protein functions (either point mutation or small insertions
CC or deletions), particularly in proteins involved in blood coagulation or
CC metabolism of fat. Protein functions which are specially detectable are
CC the Leyden mutation in factor V (associated with increased risk of deep
CC vein thrombosis), mutations in apolipoprotein (apo) genes (certain
CC alleles of apoE indicates increased risk of developing Alzheimer's
CC disease), thermostable 5,10-methylenetetrahydrofolate reductase
CC (associated with hypercysteinaemia and venous thrombosis) and factor VII
CC mutations (associated with increased risk of cardiovascular disease). The
CC method can also be applied to proteins from pathogens, e.g. viruses or
CC prions. The method does not require complex apparatus for polymerase
CC chain reactions, it is simple, standardisable and reliable and is
CC particularly suited to routine screening. It also allows mutant protein
CC in a sample to be quantified
XX
SQ Sequence 37 AA;

Query Match 86.0%; Score 43; DB 2; Length 37;
Best Local Similarity 90.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRDRGLK 10
              ||||| ||||
Db         11 TRLTRKRGLK 20
```

Search completed: December 29, 2004, 12:28:47
Job time : 63.0227 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-3
Perfect score: 48
Sequence: 1 TRLTRARGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	10	2	AAY30684	Aay30684 Apo-B100
2	45	93.8	10	2	AAY30686	Aay30686 Apo-B100
3	44	91.7	10	2	AAY30685	Aay30685 Apo-B100
4	43	89.6	10	2	AAY30682	Aay30682 Apo-B100
5	43	89.6	10	2	AAY30687	Aay30687 Apo-B100
6	43	89.6	11	2	AAW57205	Aaw57205 Apo B bin
7	43	89.6	13	2	AAW57207	Aaw57207 Apo B 100
8	43	89.6	15	2	AAW41261	Aaw41261 Apolipopr
9	43	89.6	15	2	AAW96892	Aaw96892 ApoB-100
10	43	89.6	20	6	ABJ37575	Abj37575 Heparin b
11	43	89.6	22	2	AAW57208	Aaw57208 Apo B 100
12	43	89.6	22	2	AAW57209	Aaw57209 Apo B 100
13	43	89.6	34	5	AAE14541	Aae14541 Human apo
14	43	89.6	36	2	AAW96876	Aaw96876 Nucleic a
15	43	89.6	37	2	AAW64587	Aaw64587 Human apo
16	43	89.6	51	2	AAW96845	Aaw96845 Nucleic a
17	43	89.6	343	4	ABB37687	Abb37687 Peptide #
18	43	89.6	343	4	ABG52504	Abg52504 Human liv
19	43	89.6	377	2	AAR72704	Aar72704 Human apo
20	43	89.6	377	2	AAR34031	Aar34031 Sequence
21	43	89.6	2463	8	ADJ57400	Adj57400 Human apo
22	43	89.6	3923	2	AAY31237	Aay31237 Human Apo
23	43	89.6	4536	2	AAW41262	Aaw41262 Apolipopr
24	43	89.6	4536	2	AAW96826	Aaw96826 Amino aci
25	43	89.6	4560	5	AAU98981	Aau98981 Human apo
26	43	89.6	4561	7	ADD48677	Add48677 Human Pro
27	43	89.6	4563	5	AAO15893	Aao15893 Human apo
28	43	89.6	4563	6	ABR40253	Abr40253 Human ali
29	43	89.6	4563	6	ABU79140	Abu79140 Apolipopr
30	43	89.6	4563	7	ADF43408	Adf43408 Apolipopr
31	43	89.6	4563	8	ADH18871	Adh18871 Human apo
32	43	89.6	4563	8	ADH18870	Adh18870 Human apo
33	43	89.6	4563	8	ADO33445	Ado33445 Human apo
34	43	89.6	4563	8	ADO33447	Ado33447 Human apo
35	43	89.6	4590	4	AAU33184	Aau33184 Novel hum
36	42	87.5	10	2	AAY30683	Aay30683 Apo-B100
37	38	79.2	10	2	AAY30690	Aay30690 Apo-B100
38	38	79.2	10	2	AAY30692	Aay30692 Apo-B100
39	38	79.2	10	2	AAY30688	Aay30688 Apo-B100
40	38	79.2	11	2	AAW57206	Aaw57206 Apo B 100
41	38	79.2	11	2	AAW87717	Aaw87717 Analogue
42	38	79.2	11	5	AAE21732	Aae21732 BSMR effe
43	38	79.2	11	6	ABU07938	Abu07938 Apoprotei
44	38	79.2	11	7	ADF56451	Adf56451 Human apo
45	38	79.2	12	2	AAW41260	Aaw41260 Apolipopr

ALIGNMENTS

RESULT 1
AAY30684
ID AAY30684 standard; peptide; 10 AA.
XX
AC AAY30684;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 48; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
||||||||||
Db 1 TRLTRARGLK 10

RESULT 2
AAY30686
ID AAY30686 standard; peptide; 10 AA.
XX
AC AAY30686;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 93.8%; Score 45; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.03;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
|||||:||||
Db 1 TRLTRSRGLK 10

RESULT 3
AAY30685
ID AAY30685 standard; peptide; 10 AA.
XX
AC AAY30685;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 91.7%; Score 44; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.047;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10

```
                  ||||| ||||
Db          1 TRLTRTRGLK 10


RESULT 4
AAY30682
ID   AAY30682 standard; peptide; 10 AA.
XX
AC   AAY30682;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             89.6%;  Score 43;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.076;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db          1 TRLTRERGLK 10


RESULT 5
AAY30687
ID   AAY30687 standard; peptide; 10 AA.
XX
AC   AAY30687;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             89.6%;  Score 43;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.076;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db          1 TRLTRQRGLK 10


RESULT 6
AAW57205
ID   AAW57205 standard; peptide; 11 AA.
XX
AC   AAW57205;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B binding site peptide 2.
XX
```

KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 12; Page 52; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B binding site
CC peptide which can be used as a component of a non-naturally occurring,
CC receptor-competent low density lipoprotein (LDL) particle of the present
CC invention. The LDL particle comprises at least 1 peptide component that
CC has at least 1 binding site for an apo B protein receptor and at least 1
CC lipophilic substituent. Also described in the invention are peptides
CC containing an apo B binding sequence with at least 70% identity with
CC sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC naturally occurring, receptor-competent LDL particles are useful as: (i)
CC drug-targeting vectors for delivering anticancer drugs to cancer cells
CC that express an apo B protein receptor, and (ii) additives for cell
CC culture media especially as growth supplements. Non-naturally occurring,
CC receptor-competent LDL particles do not require the complete apo B
CC sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 11 AA;

Query Match 89.6%; Score 43; DB 2; Length 11;
Best Local Similarity 90.0%; Pred. No. 0.083;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
|||||.||||
Db 2 TRLTRKRGLK 11

RESULT 7
AAW57207
ID AAW57207 standard; peptide; 13 AA.
XX
AC AAW57207;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide B.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 13 AA;

Query Match 89.6%; Score 43; DB 2; Length 13;
Best Local Similarity 90.0%; Pred. No. 0.098;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
|||||.||||
Db 3 TRLTRKRGLK 12

RESULT 8
AAW41261
ID AAW41261 standard; peptide; 15 AA.
XX
AC AAW41261;
XX
DT 19-MAY-1998 (first entry)
XX
DE Apolipoprotein B-100 fragment.
XX
KW Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW prothrombinase complex.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9743311-A1.
XX
PD 20-NOV-1997.
XX
PF 09-MAY-1997; 97WO-GB001255.
XX
PR 09-MAY-1996; 96GB-00009702.
XX

PA (UNLO) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI Bruckdorfer KR, Ettelaie C;
XX
DR WPI; 1998-008798/01.
XX
PT Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT used for treating or preventing coagulation, inhibiting angiogenesis,
PT cell differentiation and apoptosis.
XX
PS Disclosure; Page 22; 60pp; English.
XX
CC This sequence is an example of the peptide of the invention. It has the
CC formula (I), or their variants with one or more internal deletions,
CC insertions or substitutions, while retaining anti-coagulant properties of
CC apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC aa. Compositions containing the peptide are used for simultaneous,
CC separate or sequential treatment of cancer, particularly to prevent
CC metastatic spread. They are also used to inhibit thromboplastin-mediated
CC processes, specifically to prevent or reduce blood coagulation (e.g.
CC during or after surgery or in cases of heart attack, stroke etc.) and to
CC inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC which is active as such or as part of a 98-aa peptide, inhibits
CC activation of the prothrombinase complex; and prevents activation of
CC factor VII on the surface of thromboplastin and of platelets by thrombin.
CC It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC smaller than apoB-100, they act more quickly
XX
SQ Sequence 15 AA;

Query Match 89.6%; Score 43; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRARGLK 10
              ||||| ||||
Db          1 TRLTRKRGLK 10
```

RESULT 9
AAW96892
ID AAW96892 standard; peptide; 15 AA.
XX
AC AAW96892;
XX
DT 22-APR-1999 (first entry)
XX
DE ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from
CC human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC component of very-low density lipoproteins (VLDL), intermediate density
CC lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC present sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 15 AA;

Query Match 89.6%; Score 43; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRARGLK 10
              ||||| ||||
Db          6 TRLTRKRGLK 15
```

RESULT 10
ABJ37575
ID ABJ37575 standard; peptide; 20 AA.
XX
AC ABJ37575;
XX
DT 10-MAY-2003 (first entry)
XX
DE Heparin binding peptide sequence #28.
XX
KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS Unidentified.
XX
PN WO2003007689-A2.
XX
PD 30-JAN-2003.
XX
PF 22-JUL-2002; 2002WO-US023419.
XX
PR 20-JUL-2001; 2001US-0306726P.
XX
PA (ETHZ-) ETH ZUERICH.
PA (UYZU-) UNIV ZURICH.
XX
PI Hubbell JA, Schoenmakers R, Maynard HD;
XX
DR WPI; 2003-300420/29.
XX
PT Use of a ligand comprising of at least one sulfated or sulfonated amino
PT acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT retinopathy and hypoxia.
XX
PS Disclosure; Fig 2; 79pp; English.
XX
CC The invention relates to a novel ligand for binding a target biomolecule,
CC which comprises a peptide having at least one sulphated or sulphonated
CC amino acid and at least one amino acid chosen from neutral and positively
CC charged amino acids. The novel ligands can be used for the treatment of
CC e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC This sequence represents a heparin binding peptide relating to the

```
CC   invention
XX
SQ   Sequence 20 AA;

  Query Match             89.6%;  Score 43;  DB 6;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 11
AAW57208
ID   AAW57208 standard; peptide; 22 AA.
XX
AC   AAW57208;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide C.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
FT   Modified-site   22
FT                   /note= "attached to cholesterol"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             89.6%;  Score 43;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.17;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 12
AAW57209
ID   AAW57209 standard; peptide; 22 AA.
XX
AC   AAW57209;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide D.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             89.6%;  Score 43;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.17;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
```

```
Db          7 TRLTRKRGLK 16


RESULT 13
AAE14541
ID   AAE14541 standard; peptide; 34 AA.
XX
AC   AAE14541;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human apoB-100 derived peptide p62.
XX
KW   Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW   cardiovascular disease; coronary heart disease; pre-eclampsia;
KW   non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW   peptide p62.
XX
OS   Homo sapiens.
XX
PN   WO200206314-A2.
XX
PD   24-JAN-2002.
XX
PF   18-JUL-2001; 2001WO-GB003212.
XX
PR   18-JUL-2000; 2000GB-00017641.
XX
PA   (ARKT-) ARK THERAPEUTICS LTD.
XX
PI   Narvanen O,  Yla-Herttuala S;
XX
DR   WPI; 2002-179777/23.
XX
PT   New peptide useful in enzyme immunoassays for detecting oxidized low
PT   density lipoprotein which is a marker of coronary heart disease and other
PT   cardiovascular diseases, has affinity for oxidized low density
PT   lipoprotein.
XX
PS   Claim 6; Page 5; 21pp; English.
XX
CC   The invention relates to peptides having affinity for oxidised low
CC   density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC   is useful in an immunoassay to determine the presence, and optionally,
CC   the amount of antibodies in a sample, having affinity for oxLDL.
CC   Preferably immobilised peptide is useful for measuring the amount of
CC   autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC   from a patient for evaluating the risk of coronary heart diseases, other
CC   cardiovascular diseases, and several other disorders such as
CC   periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC   endothelial dysfunction. The peptide of the invention is stable, can be
CC   synthesised easily without the need to isolate proteins from a patient's
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             89.6%;  Score 43;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.26;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db         25 TRLTRKRGLK 34


RESULT 14
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Fig 12C; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 36 AA;

  Query Match             89.6%;  Score 43;  DB 2;  Length 36;
  Best Local Similarity   90.0%;  Pred. No. 0.27;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db         11 TRLTRKRGLK 20


RESULT 15
AAW64587
ID   AAW64587 standard; peptide; 37 AA.
XX
AC   AAW64587;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Human apolipoprotein peptide fragment #1.
XX
KW   Factor V; human; detection; protein function; blood coagulation; apo;
KW   fat metabolism; Leyden mutation; deep vein thrombosis; apolipoprotein;
KW   Alzheimer's disease; 5,10-methylenetetrahydrofolate reductase; prion;
KW   hypercysteinaemia; factor VII; cardiovascular disease; pathogen; virus.
XX
OS   Homo sapiens.
XX
PN   EP857973-A2.
```

XX
PD 12-AUG-1998.
XX
PF 12-JAN-1998; 98EP-00890007.
XX
PR 13-JAN-1997; 97AT-00000044.
XX
PA (IMMO) IMMUNO AG.
XX
PI Moritz B, Kiessig S, Lang H, Schenk V;
XX
DR WPI; 1998-416142/36.
XX
PT Detecting or quantifying mutant protein in presence of wild-type protein
PT comprises reaction with ligand - used to detect mutant blood coagulation
PT factors or apolipoproteins for diagnosing risk of thrombosis.
XX
PS Example 2; Page 9; 18pp; German.
XX
CC AAW64587 and AAW64588 are fragments of human apolipoprotein which are
CC used with Factor V protein fragments in a novel method to detect the
CC presence of a mutated protein in a sample that may also contain the
CC corresponding wild-type protein. The method is used to detect mutations
CC that alter protein functions (either point mutation or small insertions
CC or deletions), particularly in proteins involved in blood coagulation or
CC metabolism of fat. Protein functions which are specially detectable are
CC the Leyden mutation in factor V (associated with increased risk of deep
CC vein thrombosis), mutations in apolipoprotein (apo) genes (certain
CC alleles of apoE indicates increased risk of developing Alzheimer's
CC disease), thermostable 5,10-methylenetetrahydrofolate reductase
CC (associated with hypercysteinaemia and venous thrombosis) and factor VII
CC mutations (associated with increased risk of cardiovascular disease). The
CC method can also be applied to proteins from pathogens, e.g. viruses or
CC prions. The method does not require complex apparatus for polymerase
CC chain reactions, it is simple, standardisable and reliable and is
CC particularly suited to routine screening. It also allows mutant protein
CC in a sample to be quantified
XX
SQ Sequence 37 AA;

Query Match 89.6%; Score 43; DB 2; Length 37;
Best Local Similarity 90.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRARGLK 10
              ||||| ||||
Db         11 TRLTRKRGLK 20
```

Search completed: December 29, 2004, 12:28:48
Job time : 62.0227 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 9.65909 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-3
Perfect score: 48
Sequence: 1 TRLTRARGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	89.6	596	2	S32802	apolipoprotein B -
2	43	89.6	4563	1	LPHUB	apolipoprotein B-1
3	39	81.2	269	2	C60950	apolipoprotein B-1
4	39	81.2	779	2	JH0102	apolipoprotein B -
5	37	77.1	173	2	G87383	acetyltransferase,
6	37	77.1	275	2	E60950	apolipoprotein B-1
7	36	75.0	309	2	AH0906	conserved hypothet
8	35	72.9	208	2	E72514	hypothetical prote
9	35	72.9	388	1	DEHUPT	pyruvate dehydroge
10	34	70.8	233	1	C48560	UL56 protein - hum
11	34	70.8	309	1	E65112	hypothetical 34.6
12	34	70.8	309	2	E85985	hypothetical prote
13	34	70.8	309	2	B91140	hypothetical prote
14	34	70.8	398	2	D96795	probable DnaJ prot
15	33	68.8	171	2	S55958	probable membrane
16	33	68.8	274	2	A60950	apolipoprotein B-1
17	33	68.8	387	2	D70975	hypothetical prote
18	33	68.8	509	2	A35540	cruciferin 1 precu
19	33	68.8	542	2	D83041	probable two-compo
20	33	68.8	784	2	JH0101	apolipoprotein B-1
21	33	68.8	829	2	T32744	hypothetical prote
22	33	68.8	1049	2	T18833	Na+/K+-exchanging
23	33	68.8	4613	2	T17409	polyketide synthas
24	32	66.7	216	2	G90222	lSU ribosomal prot
25	32	66.7	225	2	H70665	hypothetical prote
26	32	66.7	232	1	S28609	phosphoadenylyl-su
27	32	66.7	272	2	E83363	hypothetical prote
28	32	66.7	290	2	S39854	traX protein - Str
29	32	66.7	307	2	AE0432	conserved hypothet
30	32	66.7	329	2	T17033	leucine rich repea
31	32	66.7	390	2	D83057	probable aminotran
32	32	66.7	393	2	S48288	probable phosphopr
33	32	66.7	425	2	F97512	clpx (AF218420) [i
34	32	66.7	425	2	AF2731	ATP-dependent Clp
35	32	66.7	442	2	F97698	probable thiophene
36	32	66.7	442	2	AG2924	tRNA modification
37	32	66.7	456	2	AE0752	flagellum-specific
38	32	66.7	456	2	C42364	H+-transporting tw
39	32	66.7	457	2	H90963	flagellum-specific
40	32	66.7	457	2	B64958	H+-transporting tw
41	32	66.7	460	2	B87455	DNA repair protein
42	32	66.7	484	2	S40051	starch synthase (E
43	32	66.7	484	2	AI0222	H+-transporting tw
44	32	66.7	490	2	H70538	probable ppdK prot
45	32	66.7	496	2	H85811	flagellum-specific

ALIGNMENTS

RESULT 1
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchi
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional r
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g9301
C;Superfamily: apolipoprotein B

Query Match 89.6%; Score 43; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 0.76;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
||||| ||||
Db 226 TRLTRKRGLK 235

RESULT 2
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A2
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Sc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNI
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: r
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986

A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732

R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.,
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than on
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place o
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

  Query Match 89.6%; Score 43; DB 1; Length 4563;
  Best Local Similarity 90.0%; Pred. No. 5.2;
  Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
 |||||.||||
Db 3385 TRLTRKRGLK 3394

RESULT 3
C60950
apolipoprotein B-100 - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: C60950
A;Molecule type: DNA
A;Residues: 1-269 <LAW>
A;Cross-references: UNIPROT:Q60537; UNIPROT:Q60536
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match 81.2%; Score 39; DB 2; Length 269;
  Best Local Similarity 80.0%; Pred. No. 2.4;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
 :||||.||||
Db 216 SRLTRKRGLK 225

RESULT 4
JH0102
apolipoprotein B - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0102
R;Smith, T.J.
submitted to GenBank, June 1990
A;Reference number: A38864
A;Accession: JH0102
A;Molecule type: DNA
A;Residues: 1-779 <SMI>
A;Cross-references: UNIPROT:Q60536; GB:M35187
A;Note: this is a revision to the sequence from reference JH0101
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Contents: annotation
A;Note: this sequence has been revised in reference A38864
C;Genetics:
A;Gene: apoB
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

  Query Match 81.2%; Score 39; DB 2; Length 779;
  Best Local Similarity 80.0%; Pred. No. 6.5;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
 :||||.||||
Db 642 SRLTRKRGLK 651

RESULT 5
G87383
acetyltransferase, GNAT family [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87383
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87383
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <STO>
A;Cross-references: UNIPROT:Q9A9B1; GB:AE005673; NID:g13422385; PIDN:AAK23067.1; GSPDB:
C;Genetics:
A;Gene: CC1083

  Query Match 77.1%; Score 37; DB 2; Length 173;
  Best Local Similarity 88.9%; Pred. No. 4;
  Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGL 9
 |||.|||||
Db 49 TRLMRARGL 57

RESULT 6
E60950
apolipoprotein B-100 - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: E60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LD

```
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: E60950
A;Molecule type: mRNA
A;Residues: 1-275 <LAW>
A;Cross-references: UNIPROT:Q7LZ77
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match             77.1%;  Score 37;  DB 2;  Length 275;
  Best Local Similarity   80.0%;  Pred. No. 6.3;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              | ||| ||||
Db        221 TSLTRKRGLK 230


RESULT 7
AH0906
conserved hypothetical protein STY3508 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0906
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0906
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07846.1; PID:g16504394; GSPDB:GN00176
C;Genetics:
A;Gene: STY3508
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match             75.0%;  Score 36;  DB 2;  Length 309;
  Best Local Similarity   70.0%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||: | ||||
Db        170 TRIARERGLK 179


RESULT 8
E72514
hypothetical protein APE2090 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E72514
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-208 <KAW>
A;Cross-references: UNIPROT:Q9YA48; DDBJ:AP000063; NID:g5105654; PIDN:BAA81101.1; PID:g5
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2090
C;Superfamily: dTMP kinase

  Query Match             72.9%;  Score 35;  DB 2;  Length 208;
  Best Local Similarity   77.8%;  Pred. No. 12;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTRARGLK 10
              || ||||:|
Db        165 RLARARGVK 173


RESULT 9
DEHUPT
pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain precursor, testis-specific -
N;Alternate names: pyruvate dehydrogenase complex, E1 component alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A37104
R;Dahl, H.H.M.; Brown, R.M.; Hutchison, W.M.; Maragos, C.; Brown, G.K.
Genomics 8, 225-232, 1990
A;Title: A testis-specific form of the human pyruvate dehydrogenase E1alpha subunit is c
A;Reference number: A37104; MUID:91065637; PMID:2249846
A;Accession: A37104
A;Molecule type: mRNA
A;Residues: 1-388 <DAH>
A;Cross-references: UNIPROT:P29803; GB:M86808; GB:J04769; NID:g190789; PIDN:AAA60232.1;
C;Genetics:
A;Gene: GDB:PDHA2
A;Cross-references: GDB:120711; OMIM:179061
A;Map position: 4q22-4q23
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin
C;Keywords: flavoprotein; heterotetramer; mitochondrion; oxidoreductase; phosphoprotein,
F;1-27/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;28-388/Product: pyruvate dehydrogenase (lipoamide) alpha chain #status predicted <MAT>
F;183-232/Domain: thiamin pyrophosphate-binding domain homology <TPB>
F;230/Binding site: phosphate (Ser) (covalent) #status predicted
F;291/Binding site: phosphate (Ser) (covalent) #status predicted
F;298/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             72.9%;  Score 35;  DB 1;  Length 388;
  Best Local Similarity   80.0%;  Pred. No. 22;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              | |||| |||
Db         52 TVLTRAEGLK 61


RESULT 10
C48560
UL56 protein - human herpesvirus 1 (strain HFEM)
C;Species: human herpesvirus 1
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: C48560
R;Rosen-Wolff, A.; Frank, S.; Raab, K.; Moyal, M.; Becker, Y.; Darai, G.
Virus Res. 25, 189-199, 1992
A;Title: Determination of the coding capacity of the BamHI DNA fragment B of apathogenic
A;Reference number: A48560; MUID:93070559; PMID:1332274
A;Accession: C48560
A;Molecule type: DNA
A;Residues: 1-233 <ROS>
A;Cross-references: UNIPROT:P36297
A;Note: sequence extracted from NCBI backbone (NCBIN:117573, NCBIP:117577)
C;Genetics:
A;Gene: UL56
C;Superfamily: herpesvirus UL56 protein

  Query Match             70.8%;  Score 34;  DB 1;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTRARG 8
              |||||||
Db        171 RLTRARG 177


RESULT 11
E65112
```

hypothetical 34.6 kD protein in arcB-gltB intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E65112
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E65112
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-309 <BLAT>
A;Cross-references: UNIPROT:P45476; GB:AE000400; GB:U00096; NID:g2367203; PIDN:AAC76243.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match 70.8%; Score 34; DB 1; Length 309;
  Best Local Similarity 70.0%; Pred. No. 29;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
|:| | ||||
Db 170 TQLARQRGLK 179

RESULT 12
E85985
hypothetical protein yhcC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85985
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85985
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <STO>
A;Cross-references: UNIPROT:P45476; GB:AE005174; NID:g12517832; PIDN:AAG58345.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match 70.8%; Score 34; DB 2; Length 309;
  Best Local Similarity 70.0%; Pred. No. 29;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
|:| | ||||
Db 170 TQLARQRGLK 179

RESULT 13
B91140
hypothetical protein ECs4090 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B91140
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91140
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <HAY>
A;Cross-references: UNIPROT:P45476; GB:BA000007; PIDN:BAB37513.1; PID:g13363563; GSPDB:(
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4090
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match 70.8%; Score 34; DB 2; Length 309;
  Best Local Similarity 70.0%; Pred. No. 29;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
|:| | ||||
Db 170 TQLARQRGLK 179

RESULT 14
D96795
probable DnaJ protein, 19794-17391 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96795
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96795
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <STO>
A;Cross-references: UNIPROT:Q9SRE3; GB:AE005173; NID:g6143904; PIDN:AAF04450.1; GSPDB:GI
C;Genetics:
A;Gene: F28O16.7
A;Map position: 1

  Query Match 70.8%; Score 34; DB 2; Length 398;
  Best Local Similarity 100.0%; Pred. No. 37;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TRARGLK 10
|||||||
Db 299 TRARGLK 305

RESULT 15
S55958
probable membrane protein YLR402w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L8084.5
C;Species: Saccharomyces cerevisiae
C;Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: S55958
R;Du, Z.
submitted to the EMBL Data Library, January 1995
A;Description: The sequence of S. cerevisiae cosmid 8084.
A;Reference number: S55944
A;Accession: S55958
A;Molecule type: DNA
A;Residues: 1-171 <DUZ>
A;Cross-references: UNIPROT:Q06057; EMBL:U19729; NID:g625097; PID:g625102; GSPDB:GN0001
A;Experimental source: strain S288C (AB972)
C;Genetics:
A;Gene: MIPS:YLR402w
A;Cross-references: SGD:S0004394
A;Map position: 12R
C;Superfamily: Saccharomyces probable membrane protein YLR402w
C;Keywords: transmembrane protein
F;24-40/Domain: transmembrane #status predicted <TM1>

F;130-146/Domain: transmembrane #status predicted <TM2>

Query Match 68.8%; Score 33; DB 2; Length 171;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRARG 8
              ||| ||||
Db         79 TRLRRARG 86
```

Search completed: December 29, 2004, 12:39:01
Job time : 11.6591 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 58.4091 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-3
Perfect score: 48
Sequence: 1 TRLTRARGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	89.6	414	2	Q7YQR5	Q7yqr5 aotus vocif
2	43	89.6	596	2	Q28473	macaca fasc
3	43	89.6	3262	2	Q13788	Q13788 homo sapien
4	43	89.6	4563	1	APB_HUMAN	P04114 homo sapien
5	43	89.6	4563	2	Q7Z600	Q7z600 homo sapien
6	39	81.2	421	2	Q7TN68	Q7tn68 glaucomys v
7	39	81.2	432	2	Q7YR10	Q7yr10 diceros bic
8	39	81.2	436	2	Q7YQM8	Q7yqm8 nyctimene a
9	39	81.2	438	2	Q7YQM7	Q7yqm7 pteropus hy
10	39	81.2	438	2	Q7YR04	Q7yr04 rousettus a
11	39	81.2	445	2	Q7YR08	Q7yr08 chaetophrac
12	39	81.2	445	2	Q7TN64	Q7tn64 agouti paca
13	39	81.2	445	2	Q7TN71	Q7tn71 hydrochoeru
14	39	81.2	445	2	Q7TN72	Q7tn72 erethizon d
15	39	81.2	780	2	Q60536	Q60536 mesocricetu
16	39	81.2	780	2	Q60537	Q60537 mesocricetu
17	38	79.2	914	2	Q6NAC6	Q6nac6 rhodopseudo
18	38	79.2	914	2	CAE26702	Cae26702 rhodopseu
19	37	77.1	173	2	Q9A9B1	Q9a9b1 caulobacter
20	37	77.1	275	2	Q7LZ77	Q7lz77 gallus gall
21	37	77.1	387	2	Q7YQN2	Q7yqn2 phalanger o
22	37	77.1	400	2	Q7YQM9	Q7yqm9 ornithorhyn
23	37	77.1	405	2	Q7YQN0	Q7yqn0 tachyglossu
24	37	77.1	445	2	Q7TN70	Q7tn70 dinomys bra
25	36	75.0	153	2	Q9FXM2	Q9fxm2 arabidopsis
26	36	75.0	202	2	Q8L8T0	Q8l8t0 arabidopsis
27	36	75.0	202	2	Q9LVA4	Q9lva4 arabidopsis
28	36	75.0	309	2	Q8XFV9	Q8xfv9 salmonella
29	36	75.0	309	2	Q7CPN5	Q7cpn5 salmonella
30	36	75.0	407	2	Q7TN65	Q7tn65 atherurus a
31	36	75.0	412	2	Q7TN69	Q7tn69 hystrix bra
32	36	75.0	2073	2	Q73YT8	Q73yt8 mycobacteri
33	36	75.0	2073	2	AAS04184	Aas04184 mycobacte
34	35	72.9	158	2	Q8IJ19	Q8ij19 plasmodium
35	35	72.9	208	1	KTHY_AERPE	Q9ya48 aeropyrum p
36	35	72.9	289	2	Q8DUW3	Q8duw3 streptococc
37	35	72.9	365	2	Q746G7	Q746g7 thermus the
38	35	72.9	365	2	AAS82418	Aas82418 thermus t
39	35	72.9	388	1	ODPT_HUMAN	P29803 homo sapien
40	35	72.9	403	2	Q6NXQ1	Q6nxq1 homo sapien
41	35	72.9	403	2	AAH66953	Aah66953 homo sapi
42	35	72.9	566	2	Q8IIX8	Q8iix8 plasmodium
43	35	72.9	600	2	Q8FRZ6	Q8frz6 corynebacte
44	35	72.9	675	2	Q8IAQ5	Q8iaq5 plasmodium
45	34	70.8	179	2	Q98FE8	Q98fe8 rhizobium l

ALIGNMENTS

RESULT 1
Q7YQR5
ID Q7YQR5 PRELIMINARY; PRT; 414 AA.
AC Q7YQR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein B 100 (Fragment).
GN Name=apoB-100;
OS Aotus vociferans (Spix's owl monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=57176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22761261; PubMed=12878460;
RA Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT "A new phylogenetic marker, apolipoprotein B, provides compelling
RT evidence for eutherian relationships.";
RL Mol. Phylogenet. Evol. 28:225-240(2003).
DR EMBL; AF548396; AAP97352.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45955 MW; EEFA8492157E1BDE CRC64;

Query Match 89.6%; Score 43; DB 2; Length 414;
Best Local Similarity 90.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRARGLK 10
||||| ||||
Db 258 TRLTRKRGLK 267

RESULT 2
Q28473
ID Q28473 PRELIMINARY; PRT; 596 AA.
AC Q28473;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Apolipoprotein B (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92075708; PubMed=1742325;
RA Pape M.E., Castle C.K., Murray R.W., Funk G.M., Hunt C.E.,
RA Marotti K.R., Melchior G.W.;

```
RT   "Apo B metabolism in the cynomolgus monkey: evidence for post-
RT   transcriptional regulation.";
RL   Biochim. Biophys. Acta 1086:326-334(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Murray R.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X15737; CAA33755.1; -.
DR   PIR; S32802; S32802.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     596    596
SQ   SEQUENCE   596 AA;  66757 MW;  B13BBA74E25C3120 CRC64;

  Query Match             89.6%;  Score 43;  DB 2;  Length 596;
  Best Local Similarity   90.0%;  Pred. No. 4.3;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db        226 TRLTRKRGLK 235


RESULT 3
Q13788
ID   Q13788      PRELIMINARY;      PRT;  3262 AA.
AC   Q13788;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   APOB protein (Fragment).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191999; PubMed=2883086;
RA   Carlsson P., Darnfors C., Olofsson S.O., Bjursell G.;
RT   "Analysis of the human apolipoprotein B gene; complete structure of
RT   the B-74 region.";
RL   Gene 49:29-51(1986).
DR   EMBL; M15421; AAA51758.1; -.
DR   PIR; A27850; LPHUB.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005319; F:lipid transporter activity; NAS.
DR   GO; GO:0006869; P:lipid transport; NAS.
FT   NON_TER       1      1
SQ   SEQUENCE   3262 AA;  370140 MW;  56603BC0618DD40D CRC64;

  Query Match             89.6%;  Score 43;  DB 2;  Length 3262;
  Best Local Similarity   90.0%;  Pred. No. 25;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db       2084 TRLTRKRGLK 2093


RESULT 4
APB_HUMAN
ID   APB_HUMAN      STANDARD;      PRT;  4563 AA.
AC   P04114; O00502; Q13787;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein
DE   B-48 (Apo B-48)].
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87016385; PubMed=3763409;
RA   Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.,
RA   Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
RT   "Complete cDNA and derived protein sequence of human apolipoprotein B-
RT   100.";
RL   Nucleic Acids Res. 14:7501-7503(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=88003974; PubMed=3652907;
RA   Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
RA   Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
RT   "DNA sequence of the human apolipoprotein B gene.";
RL   DNA 6:363-372(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS ILE-98 AND GLU-4181.
RX   MEDLINE=87008488; PubMed=3759943;
RA   Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
RA   Gotto A.M. Jr., Chan L.;
RT   "The complete cDNA and amino acid sequence of human apolipoprotein B-
RT   100.";
RL   J. Biol. Chem. 261:12918-12921(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87041416; PubMed=3464946;
RA   Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
RA   Lee N., Brewer H.B. Jr.;
RT   "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
RT   derived amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87161758; PubMed=3030729;
RA   Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
RA   Zannis V.I.;
RT   "The complete sequence and structural analysis of human apolipoprotein
RT   B-100: relationship between apoB-100 and apoB-48 forms.";
RL   EMBO J. 5:3495-3507(1986).
RN   [6]
RP   SEQUENCE OF 709-906 FROM N.A.
RX   MEDLINE=85270450; PubMed=3860836;
RA   Deeb S.S., Motulsky A.G., Albers J.J.;
RT   "A partial cDNA clone for human apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
RN   [7]
RP   SEQUENCE OF 3056-3159 FROM N.A.
RX   MEDLINE=86041888; PubMed=3903660;
RA   Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
RA   Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
RT   "Human apolipoprotein B: identification of cDNA clones and
RT   characterization of mRNA.";
RL   Nucleic Acids Res. 13:6937-6953(1985).
RN   [8]
RP   SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=86093680; PubMed=3841204;
RA   Carlsson P., Olofsson S.O., Bondjers G., Darnfors C., Wiklund O.,
RA   Bjursell G.;
RT   "Molecular cloning of human apolipoprotein B cDNA.";
RL   Nucleic Acids Res. 13:8813-8826(1985).
RN   [9]
RP   SEQUENCE OF 3109-4563 FROM N.A.
RX   MEDLINE=85300528; PubMed=2994225;
RA   Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F., Urdea M.S.,
RA   Levy-Wilson B., Powell L.M., Pease R.J., Eddy R., Nakai H., Byers M.,
RA   Priestley L.M., Robertson E., Rall L.B., Betsholtz C., Shows T.B.,
RA   Mahley R.W., Scott J.;
RT   "Human apolipoprotein B: structure of carboxyl-terminal domains, sites
RT   of gene expression, and chromosomal localization.";
RL   Science 230:37-43(1985).
RN   [10]
```

```
RP   SEQUENCE OF 1-291 FROM N.A.
RX   MEDLINE=86149325; PubMed=3513177;
RA   Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
RA   Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
RT   "Isolation of a cDNA clone encoding the amino-terminal region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
RN   [11]
RP   SEQUENCE OF 1-1670 FROM N.A., AND VARIANT ILE-98.
RX   MEDLINE=86287319; PubMed=3461454;
RA   Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W., Yamanaka M.,
RA   Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
RT   "Analysis of cDNA clones encoding the entire B-26 region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
RN   [12]
RP   PARTIAL SEQUENCE, AND IDENTIFICATION OF APO-B48.
RX   MEDLINE=88018019; PubMed=3659919;
RA   Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
RA   Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
RA   Gotto A.M. Jr., Li W.-H., Chan L.;
RT   "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
RT   specific in-frame stop codon.";
RL   Science 238:363-366(1987).
RN   [13]
RP   DOMAINS.
RX   MEDLINE=87039351; PubMed=3773997;
RA   Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
RA   Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
RA   Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
RA   Levy-Wilson B., Scott J.;
RT   "Complete protein sequence and identification of structural domains of
RT   human apolipoprotein B.";
RL   Nature 323:734-738(1986).
RN   [14]
RP   DOMAINS.
RA   Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
RA   Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
RA   Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
RT   "Sequence, structure, receptor-binding domains and internal repeats of
RT   human apolipoprotein B-100.";
RL   Nature 323:738-742(1986).
RN   [15]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=86242245; PubMed=3087360;
RA   Dashti N., Lee D.M., Mok T.;
RT   "Apolipoprotein B is a calcium binding protein.";
RL   Biochem. Biophys. Res. Commun. 137:493-499(1986).
RN   [16]
RP   PALMITOYLATION OF CYS-1112.
RX   MEDLINE=20143590; PubMed=10679026;
RA   Zhao Y., McCabe J.B., Vance J., Berthiaume L.G.;
RT   "Palmitoylation of apolipoprotein B is required for proper
RT   intracellular sorting and transport of cholesteroyl esters and
RT   triglycerides.";
RL   Mol. Biol. Cell 11:721-734(2000).
RN   [17]
RP   VARIANT SER-4338.
RX   MEDLINE=91071750; PubMed=1979313;
RA   Navajas M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
RA   Cunny G., Cambien F., Roizes G.;
RT   "Detection by denaturing gradient gel electrophoresis of a new
RT   polymorphism in the apolipoprotein B gene.";
RL   Hum. Genet. 86:91-93(1990).
RN   [18]
RP   VARIANT FDB GLN-3527.
RX   MEDLINE=89098975; PubMed=2563166;
RA   Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
RA   McCarthy B.J.;
RT   "Association between a specific apolipoprotein B mutation and familial
RT   defective apolipoprotein B-100.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
RN   [19]
RP   VARIANT LEU-2739.
RX   MEDLINE=91016974; PubMed=2216805;
RA   Huang L.-S., Gavish D., Breslow J.L.;
RT   "Sequence polymorphism in the human apoB gene at position 8344.";
RL   Nucleic Acids Res. 18:5922-5922(1990).
RN   [20]
RP   VARIANT FDB CYS-3558.
RX   MEDLINE=95190020; PubMed=7883971;
RA   Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
RA   Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
RT   "Familial ligand-defective apolipoprotein B. Identification of a new
RT   mutation that decreases LDL receptor binding affinity.";
RL   J. Clin. Invest. 95:1225-1234(1995).
RN   [21]
RP   VARIANTS LEU-1437; SER-1914; LYS-2566; THR-3121; ALA-3945; MET-4128
RP   AND THR-4481.
RX   MEDLINE=97044521; PubMed=8889592;
RA   Poirier O., Ricard S., Behague I., Souriau C., Evans A.E.,
RA   Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
RT   "Detection of new variants in the apolipoprotein B (Apo B) gene by
RT   PCR-SSCP.";
RL   Hum. Mutat. 8:282-285(1996).
RN   [22]
RP   VARIANTS FDB GLN-3527 AND CYS-3558.
RX   MEDLINE=97403938; PubMed=9259199;
RA   Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
RA   Krempf M., Giraudet P., Junien C., Boileau C.;
RT   "Familial ligand-defective apolipoprotein B-100: simultaneous
RT   detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French
RT   population.";
RL   Hum. Mutat. 10:160-163(1997).
RN   [23]
RP   VARIANTS SER-1914; ARG-1923; LEU-2739; ASP-3319; THR-3427; GLN-3432
RP   AND ILE-3921.
RX   MEDLINE=98141125; PubMed=9490296;
RA   Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
RT   "Screening for mutations of the apolipoprotein B gene causing
RT   hypocholesterolemia.";
RL   Hum. Genet. 102:44-49(1998).
CC   -!- FUNCTION: Apolipoprotein B is a major protein constituent of
CC       chylomicrons (apo B-48), LDL (apo B-100) and VLDL (apo B-100). Apo
CC       B-100 functions as a recognition signal for the cellular binding
CC       and internalization of LDL particles by the apoB/E receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.

  Query Match             89.6%;  Score 43;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 36;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 5
Q7Z600
ID   Q7Z600      PRELIMINARY;      PRT;  4563 AA.
AC   Q7Z600;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Including Ag(X) antigen).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY324608; AAP72970.1; -.
DR   GO; GO:0005319; F:lipid transporter activity; IEA.
DR   GO; GO:0006869; P:lipid transport; IEA.
DR   InterPro; IPR009454; DUF1081.
DR   InterPro; IPR001747; Lipid_transprt_N.
DR   Pfam; PF06448; DUF1081; 1.
DR   Pfam; PF01347; Vitellogenin_N; 1.
DR   SMART; SM00638; LPD_N; 1.
KW   Lipoprotein.
SQ   SEQUENCE   4563 AA;  515553 MW;  030B34167CEDC63C CRC64;

  Query Match             89.6%;  Score 43;  DB 2;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 36;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 6
Q7TN68
ID   Q7TN68      PRELIMINARY;      PRT;   421 AA.
AC   Q7TN68;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Glaucomys volans (Southern flying squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Petauristinae;
OC   Glaucomys.
OX   NCBI_TaxID=64683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243379; AAP50767.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     421    421
SQ   SEQUENCE   421 AA;  46747 MW;  D47B77BD4F864FD1 CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 421;
  Best Local Similarity   80.0%;  Pred. No. 20;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db       264 SRLTRKRGLK 273


RESULT 7
Q7YR10
ID   Q7YR10      PRELIMINARY;      PRT;   432 AA.
AC   Q7YR10;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243375; AAP50763.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     432    432
SQ   SEQUENCE   432 AA;  48171 MW;  F27B7AB39604732C CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 432;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db       275 SRLTRKRGLK 284


RESULT 8
Q7YQM8
ID   Q7YQM8      PRELIMINARY;      PRT;   436 AA.
AC   Q7YQM8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Nyctimene albiventer (Common tube-nosed fruit bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Nyctimene.
OX   NCBI_TaxID=48988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548435; AAP97391.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     436    436
SQ   SEQUENCE   436 AA;  48717 MW;  1C4A7EAD72D2C629 CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 436;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db       279 SRLTRKRGLK 288


RESULT 9
Q7YQM7
ID   Q7YQM7      PRELIMINARY;      PRT;   438 AA.
AC   Q7YQM7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Pteropus hypomelanus (Small flying fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Pteropus.
OX   NCBI_TaxID=9405;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
```

```
DR   EMBL; AF548436; AAP97392.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48734 MW;  2BD85BCBF4E2CC41 CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        281 SRLTRKRGLK 290


RESULT 10
Q7YR04
ID   Q7YR04      PRELIMINARY;      PRT;   438 AA.
AC   Q7YR04;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Rousettus amplexicaudatus (Common rousette).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Rousettus.
OX   NCBI_TaxID=58083;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243383; AAP50771.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48597 MW;  41C890DEAF95C872 CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        281 SRLTRKRGLK 290


RESULT 11
Q7YR08
ID   Q7YR08      PRELIMINARY;      PRT;   445 AA.
AC   Q7YR08;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Chaetophractus villosus (South American armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Chaetophractus.
OX   NCBI_TaxID=29080;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243378; AAP50766.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49564 MW;  2DA5DC3ED2F0FDD2 CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 12
Q7TN64
ID   Q7TN64      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN64;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Agouti paca (Paca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Agoutidae; Agouti.
OX   NCBI_TaxID=108852;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548417; AAP97373.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49721 MW;  34AF7ABE90F121EF CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 13
Q7TN71
ID   Q7TN71      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Hydrochoerus hydrochaeris (Capybara) (Carpincho).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Hydrochaeridae;
OC   Hydrochaeris.
OX   NCBI_TaxID=10149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243369; AAP50757.1; -.
DR   InterPro; IPR000871; Beta_lactamase_A.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49520 MW;  CBBA2DD53D7A18D2 CRC64;
```

```
  Query Match             81.2%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 14
Q7TN72
ID   Q7TN72      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN72;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Erethizon dorsatum (North American porcupine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Erethizontidae;
OC   Erethizon.
OX   NCBI_TaxID=34844;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243368; AAP50756.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49617 MW;  9572FE5F5E7625F2 CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 15
Q60536
ID   Q60536      PRELIMINARY;      PRT;   780 AA.
AC   Q60536;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hamster apolipoprotein (apoB) (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236327; PubMed=2332175;
RA   Smith T.J., Hautamaa D., Maeda N.;
RT   "Sequence of the putative low-density lipoprotein receptor-binding
RT   regions of apolipoprotein B in mouse and hamster.";
RL   Gene 87:309-310(1990).
DR   EMBL; M35187; AAA37059.1; -.
DR   PIR; C60950; C60950.
DR   PIR; JH0102; JH0102.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     780    780
SQ   SEQUENCE   780 AA;  86625 MW;  E371D1B2079D8F7E CRC64;

  Query Match             81.2%;  Score 39;  DB 2;  Length 780;
  Best Local Similarity   80.0%;  Pred. No. 38;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRARGLK 10
              :|||| ||||
Db        642 SRLTRKRGLK 651


Search completed: December 29, 2004, 12:37:30
Job time : 59.5202 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-4
Perfect score: 49
Sequence: 1 TRLTRTRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	49	100.0	10	2	AAY30685	Aay30685 Apo-B100
2	45	91.8	10	2	AAY30686	Aay30686 Apo-B100
3	44	89.8	10	2	AAY30684	Aay30684 Apo-B100
4	43	87.8	10	2	AAY30683	Aay30683 Apo-B100
5	43	87.8	10	2	AAY30682	Aay30682 Apo-B100
6	43	87.8	10	2	AAY30687	Aay30687 Apo-B100
7	43	87.8	11	2	AAW57205	Aaw57205 Apo B bin
8	43	87.8	13	2	AAW57207	Aaw57207 Apo B 100
9	43	87.8	15	2	AAW41261	Aaw41261 Apolipopr
10	43	87.8	15	2	AAW96892	Aaw96892 ApoB-100
11	43	87.8	20	6	ABJ37575	Abj37575 Heparin b
12	43	87.8	22	2	AAW57208	Aaw57208 Apo B 100
13	43	87.8	22	2	AAW57209	Aaw57209 Apo B 100
14	43	87.8	34	5	AAE14541	Aae14541 Human apo
15	43	87.8	36	2	AAW96876	Aaw96876 Nucleic a
16	43	87.8	37	2	AAW64587	Aaw64587 Human apo
17	43	87.8	51	2	AAW96845	Aaw96845 Nucleic a
18	43	87.8	343	4	ABB37687	Abb37687 Peptide #
19	43	87.8	343	4	ABG52504	Abg52504 Human liv
20	43	87.8	377	2	AAR72704	Aar72704 Human apo
21	43	87.8	377	2	AAR34031	Aar34031 Sequence
22	43	87.8	2463	8	ADJ57400	Adj57400 Human apo
23	43	87.8	3923	2	AAY31237	Aay31237 Human Apo
24	43	87.8	4536	2	AAW41262	Aaw41262 Apolipopr
25	43	87.8	4536	2	AAW96826	Aaw96826 Amino aci
26	43	87.8	4560	5	AAU98981	Aau98981 Human apo
27	43	87.8	4561	7	ADD48677	Add48677 Human Pro
28	43	87.8	4563	5	AAO15893	Aao15893 Human apo
29	43	87.8	4563	6	ABR40253	Abr40253 Human ali
30	43	87.8	4563	6	ABU79140	Abu79140 Apolipopr
31	43	87.8	4563	7	ADF43408	Adf43408 Apolipopr
32	43	87.8	4563	8	ADH18871	Adh18871 Human apo
33	43	87.8	4563	8	ADH18870	Adh18870 Human apo
34	43	87.8	4563	8	ADO33445	Ado33445 Human apo
35	43	87.8	4563	8	ADO33447	Ado33447 Human apo
36	43	87.8	4590	4	AAU33184	Aau33184 Novel hum
37	38	77.6	10	2	AAY30690	Aay30690 Apo-B100
38	38	77.6	10	2	AAY30692	Aay30692 Apo-B100
39	38	77.6	10	2	AAY30688	Aay30688 Apo-B100
40	38	77.6	11	2	AAW57206	Aaw57206 Apo B 100
41	38	77.6	11	2	AAW87717	Aaw87717 Analogue
42	38	77.6	11	5	AAE21732	Aae21732 BSMR effe
43	38	77.6	11	6	ABU07938	Abu07938 Apoprotei
44	38	77.6	11	7	ADF56451	Adf56451 Human apo
45	38	77.6	12	2	AAW41260	Aaw41260 Apolipopr

ALIGNMENTS

RESULT 1
AAY30685
ID AAY30685 standard; peptide; 10 AA.
XX
AC AAY30685;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
||||||||||
Db 1 TRLTRTRGLK 10

RESULT 2
AAY30686
ID AAY30686 standard; peptide; 10 AA.
XX
AC AAY30686;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 91.8%; Score 45; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.028;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
|||||:||||
Db 1 TRLTRSRGLK 10

RESULT 3
AAY30684
ID AAY30684 standard; peptide; 10 AA.
XX
AC AAY30684;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.045;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10

```
                  ||||| ||||
Db          1 TRLTRARGLK 10


RESULT 4
AAY30683
ID   AAY30683 standard; peptide; 10 AA.
XX
AC   AAY30683;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.071;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
                ||||| ||||
Db          1 TRLTRDRGLK 10


RESULT 5
AAY30682
ID   AAY30682 standard; peptide; 10 AA.
XX
AC   AAY30682;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.071;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
                ||||| ||||
Db          1 TRLTRERGLK 10


RESULT 6
AAY30687
ID   AAY30687 standard; peptide; 10 AA.
XX
AC   AAY30687;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
```

```
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.071;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db          1 TRLTRQRGLK 10


RESULT 7
AAW57205
ID   AAW57205 standard; peptide; 11 AA.
XX
AC   AAW57205;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B binding site peptide 2.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 12; Page 52; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B binding site
CC   peptide which can be used as a component of a non-naturally occurring,
CC   receptor-competent low density lipoprotein (LDL) particle of the present
CC   invention. The LDL particle comprises at least 1 peptide component that
CC   has at least 1 binding site for an apo B protein receptor and at least 1
CC   lipophilic substituent. Also described in the invention are peptides
CC   containing an apo B binding sequence with at least 70% identity with
CC   sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC   naturally occurring, receptor-competent LDL particles are useful as: (i)
CC   drug-targeting vectors for delivering anticancer drugs to cancer cells
CC   that express an apo B protein receptor, and (ii) additives for cell
CC   culture media especially as growth supplements. Non-naturally occurring,
CC   receptor-competent LDL particles do not require the complete apo B
CC   sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 11 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 11;
  Best Local Similarity   90.0%;  Pred. No. 0.079;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db          2 TRLTRKRGLK 11


RESULT 8
AAW57207
ID   AAW57207 standard; peptide; 13 AA.
XX
AC   AAW57207;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide B.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
```

XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 13 AA;

Query Match 87.8%; Score 43; DB 2; Length 13;
Best Local Similarity 90.0%; Pred. No. 0.093;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
||||| ||||
Db 3 TRLTRKRGLK 12

RESULT 9
AAW41261
ID AAW41261 standard; peptide; 15 AA.
XX
AC AAW41261;
XX
DT 19-MAY-1998 (first entry)
XX
DE Apolipoprotein B-100 fragment.
XX
KW Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW prothrombinase complex.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9743311-A1.
XX
PD 20-NOV-1997.
XX
PF 09-MAY-1997; 97WO-GB001255.
XX
PR 09-MAY-1996; 96GB-00009702.
XX
PA (UNLO) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI Bruckdorfer KR, Ettelaie C;
XX
DR WPI; 1998-008798/01.
XX
PT Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT used for treating or preventing coagulation, inhibiting angiogenesis,
PT cell differentiation and apoptosis.
XX
PS Disclosure; Page 22; 60pp; English.
XX
CC This sequence is an example of the peptide of the invention. It has the
CC formula (I), or their variants with one or more internal deletions,
CC insertions or substitutions, while retaining anti-coagulant properties of
CC apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC aa. Compositions containing the peptide are used for simultaneous,
CC separate or sequential treatment of cancer, particularly to prevent
CC metastatic spread. They are also used to inhibit thromboplastin-mediated
CC processes, specifically to prevent or reduce blood coagulation (e.g.
CC during or after surgery or in cases of heart attack, stroke etc.) and to
CC inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC which is active as such or as part of a 98-aa peptide, inhibits
CC activation of the prothrombinase complex; and prevents activation of
CC factor VII on the surface of thromboplastin and of platelets by thrombin.
CC It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC smaller than apoB-100, they act more quickly
XX
SQ Sequence 15 AA;

Query Match 87.8%; Score 43; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
||||| ||||
Db 1 TRLTRKRGLK 10

RESULT 10
AAW96892
ID AAW96892 standard; peptide; 15 AA.
XX
AC AAW96892;
XX
DT 22-APR-1999 (first entry)
XX
DE ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from

```
CC   human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC   component of very-low density lipoproteins (VLDL), intermediate density
CC   lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC   present sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 15 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 15;
  Best Local Similarity   90.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db          6 TRLTRKRGLK 15


RESULT 11
ABJ37575
ID   ABJ37575 standard; peptide; 20 AA.
XX
AC   ABJ37575;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Heparin binding peptide sequence #28.
XX
KW   Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW   rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS   Unidentified.
XX
PN   WO2003007689-A2.
XX
PD   30-JAN-2003.
XX
PF   22-JUL-2002; 2002WO-US023419.
XX
PR   20-JUL-2001; 2001US-0306726P.
XX
PA   (ETHZ-) ETH ZUERICH.
PA   (UYZU-) UNIV ZURICH.
XX
PI   Hubbell JA,  Schoenmakers R,  Maynard HD;
XX
DR   WPI; 2003-300420/29.
XX
PT   Use of a ligand comprising of at least one sulfated or sulfonated amino
PT   acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT   retinopathy and hypoxia.
XX
PS   Disclosure; Fig 2; 79pp; English.
XX
CC   The invention relates to a novel ligand for binding a target biomolecule,
CC   which comprises a peptide having at least one sulphated or sulphonated
CC   amino acid and at least one amino acid chosen from neutral and positively
CC   charged amino acids. The novel ligands can be used for the treatment of
CC   e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC   This sequence represents a heparin binding peptide relating to the
CC   invention
XX
SQ   Sequence 20 AA;

  Query Match             87.8%;  Score 43;  DB 6;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 12
AAW57208
ID   AAW57208 standard; peptide; 22 AA.
XX
AC   AAW57208;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide C.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
FT   Modified-site   22
FT                   /note= "attached to cholesterol"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.16;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16
```

```
RESULT 13
AAW57209
ID   AAW57209 standard; peptide; 22 AA.
XX
AC   AAW57209;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide D.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             87.8%;  Score 43;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.16;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 14
AAE14541
ID   AAE14541 standard; peptide; 34 AA.
XX
AC   AAE14541;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human apoB-100 derived peptide p62.
XX
KW   Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW   cardiovascular disease; coronary heart disease; pre-eclampsia;
KW   non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW   peptide p62.
XX
OS   Homo sapiens.
XX
PN   WO200206314-A2.
XX
PD   24-JAN-2002.
XX
PF   18-JUL-2001; 2001WO-GB003212.
XX
PR   18-JUL-2000; 2000GB-00017641.
XX
PA   (ARKT-) ARK THERAPEUTICS LTD.
XX
PI   Narvanen O,  Yla-Herttuala S;
XX
DR   WPI; 2002-179777/23.
XX
PT   New peptide useful in enzyme immunoassays for detecting oxidized low
PT   density lipoprotein which is a marker of coronary heart disease and other
PT   cardiovascular diseases, has affinity for oxidized low density
PT   lipoprotein.
XX
PS   Claim 6; Page 5; 21pp; English.
XX
CC   The invention relates to peptides having affinity for oxidised low
CC   density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC   is useful in an immunoassay to determine the presence, and optionally,
CC   the amount of antibodies in a sample, having affinity for oxLDL.
CC   Preferably immobilised peptide is useful for measuring the amount of
CC   autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC   from a patient for evaluating the risk of coronary heart diseases, other
CC   cardiovascular diseases, and several other disorders such as
CC   periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC   endothelial dysfunction. The peptide of the invention is stable, can be
CC   synthesised easily without the need to isolate proteins from a patient's
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             87.8%;  Score 43;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db         25 TRLTRKRGLK 34


RESULT 15
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
```

XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 16; Fig 12C; 293pp; English.
XX
CC AAW96827-77 represent nucleic acid binding domains derived from human
CC apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 36 AA;

Query Match 87.8%; Score 43; DB 2; Length 36;
Best Local Similarity 90.0%; Pred. No. 0.26;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db         11 TRLTRKRGLK 20
```

Search completed: December 29, 2004, 12:28:48
Job time : 61.0227 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 9.65909 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-4
Perfect score: 49
Sequence: 1 TRLTRTRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	87.8	596	2	S32802	apolipoprotein B -
2	43	87.8	4563	1	LPHUB	apolipoprotein B-1
3	39	79.6	269	2	C60950	apolipoprotein B-1
4	39	79.6	779	2	JH0102	apolipoprotein B -
5	37	75.5	275	2	E60950	apolipoprotein B-1
6	36	73.5	309	2	AH0906	conserved hypothet
7	35	71.4	909	2	A75337	exonuclease SbcC -
8	34	69.4	309	1	E65112	hypothetical 34.6
9	34	69.4	309	2	E85985	hypothetical prote
10	34	69.4	309	2	B91140	hypothetical prote
11	33	67.3	173	2	G87383	acetyltransferase,
12	33	67.3	274	2	A60950	apolipoprotein B-1
13	33	67.3	436	2	D70375	transcription term
14	33	67.3	449	2	T49646	hsp70 related prot
15	33	67.3	784	2	JH0101	apolipoprotein B-1
16	32	65.3	38	2	G71305	probable ribosomal
17	32	65.3	126	2	T16727	hypothetical prote
18	32	65.3	219	2	G95913	probable cell surf
19	32	65.3	232	1	S28609	phosphoadenylyl-su
20	32	65.3	272	2	E83363	hypothetical prote
21	32	65.3	290	2	S39854	traX protein - Str
22	32	65.3	329	2	T17033	leucine rich repea
23	32	65.3	393	2	S48288	probable phosphopr
24	32	65.3	412	2	T09313	immediate-early pr
25	32	65.3	460	2	B87455	DNA repair protein
26	32	65.3	484	2	S40051	starch synthase (E
27	32	65.3	670	2	G64921	probable membrane
28	32	65.3	911	2	A39967	inter-alpha-trypsi
29	32	65.3	1224	2	S73171	DNA-directed RNA p
30	31	63.3	106	2	A71072	hypothetical prote
31	31	63.3	146	2	T14681	myc-like regulator
32	31	63.3	202	2	T05763	hypothetical prote
33	31	63.3	208	2	E72514	hypothetical prote
34	31	63.3	256	2	T15383	hypothetical prote
35	31	63.3	260	2	A36949	28.9K basic DNA-bi
36	31	63.3	302	2	D83958	DNA processing pro
37	31	63.3	304	2	A98146	probable threonin
38	31	63.3	336	2	AC3142	threonine dehydrat
39	31	63.3	339	2	S62596	ubiquinol-cytochro
40	31	63.3	377	2	B90448	conserved hypothet
41	31	63.3	388	1	DEHUPT	pyruvate dehydroge
42	31	63.3	391	2	S22579	translation initia
43	31	63.3	412	2	E83061	hypothetical prote
44	31	63.3	506	2	AD3338	cobyric acid synth
45	31	63.3	614	1	S75294	ferrous iron trans

ALIGNMENTS

RESULT 1
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchi
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional r
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g9301
C;Superfamily: apolipoprotein B

Query Match 87.8%; Score 43; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db        226 TRLTRKRGLK 235
```

RESULT 2
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A2
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Sc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNI
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: r
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986

A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732

R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.,
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than on
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place o
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

```
A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

  Query Match             87.8%;  Score 43;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 5.8;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 3
C60950
apolipoprotein B-100 - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: C60950
A;Molecule type: DNA
A;Residues: 1-269 <LAW>
A;Cross-references: UNIPROT:Q60537; UNIPROT:Q60536
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match             79.6%;  Score 39;  DB 2;  Length 269;
  Best Local Similarity   80.0%;  Pred. No. 2.5;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        216 SRLTRKRGLK 225


RESULT 4
JH0102
apolipoprotein B - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0102
R;Smith, T.J.
submitted to GenBank, June 1990
A;Reference number: A38864
A;Accession: JH0102
A;Molecule type: DNA
A;Residues: 1-779 <SMI>
A;Cross-references: UNIPROT:Q60536; GB:M35187
A;Note: this is a revision to the sequence from reference JH0101
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Contents: annotation
A;Note: this sequence has been revised in reference A38864
C;Genetics:
A;Gene: apoB
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein,
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

  Query Match             79.6%;  Score 39;  DB 2;  Length 779;
  Best Local Similarity   80.0%;  Pred. No. 6.9;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        642 SRLTRKRGLK 651


RESULT 5
E60950
apolipoprotein B-100 - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: E60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDI
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: E60950
A;Molecule type: mRNA
A;Residues: 1-275 <LAW>
A;Cross-references: UNIPROT:Q7LZ77
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein,

  Query Match             75.5%;  Score 37;  DB 2;  Length 275;
  Best Local Similarity   80.0%;  Pred. No. 6.4;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              | ||| ||||
Db        221 TSLTRKRGLK 230


RESULT 6
AH0906
conserved hypothetical protein STY3508 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0906
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
```

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0906
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07846.1; PID:g16504394; GSPDB:GN00176
C;Genetics:
A;Gene: STY3508
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 73.5%; Score 36; DB 2; Length 309;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
||: | ||||
Db 170 TRIARERGLK 179

RESULT 7
A75337
exonuclease SbcC - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75337
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75337
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-909 <WHI>
A;Cross-references: UNIPROT:Q9RT44; GB:AE002032; GB:AE000513; NID:g6459715; PIDN:AAF1147
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1922
A;Map position: 1

Query Match 71.4%; Score 35; DB 2; Length 909;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RLTRTRGLK 10
|:||||| |
Db 86 RVTRTRGRK 94

RESULT 8
E65112
hypothetical 34.6 kD protein in arcB-gltB intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E65112
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E65112
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-309 <BLAT>
A;Cross-references: UNIPROT:P45476; GB:AE000400; GB:U00096; NID:g2367203; PIDN:AAC76243.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 69.4%; Score 34; DB 1; Length 309;
Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
|:| | ||||
Db 170 TQLARQRGLK 179

RESULT 9
E85985
hypothetical protein yhcC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85985
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85985
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <STO>
A;Cross-references: UNIPROT:P45476; GB:AE005174; NID:g12517832; PIDN:AAG58345.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 69.4%; Score 34; DB 2; Length 309;
Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
|:| | ||||
Db 170 TQLARQRGLK 179

RESULT 10
B91140
hypothetical protein ECs4090 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B91140
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91140
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <HAY>
A;Cross-references: UNIPROT:P45476; GB:BA000007; PIDN:BAB37513.1; PID:g13363563; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4090
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 69.4%; Score 34; DB 2; Length 309;
Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
|:| | ||||
Db 170 TQLARQRGLK 179

RESULT 11
G87383
acetyltransferase, GNAT family [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

```
C;Accession: G87383
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87383
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <STO>
A;Cross-references: UNIPROT:Q9A9B1; GB:AE005673; NID:g13422385; PIDN:AAK23067.1; GSPDB:G
C;Genetics:
A;Gene: CC1083

  Query Match             67.3%;  Score 33;  DB 2;  Length 173;
  Best Local Similarity   77.8%;  Pred. No. 27;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGL 9
              ||| | |||
Db         49 TRLMRARGL 57


RESULT 12
A60950
apolipoprotein B-100 - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: A60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: A60950
A;Molecule type: mRNA
A;Residues: 1-274 <LAW>
A;Cross-references: UNIPROT:Q7M2U9
A;Note: authors translated the codon GAT for residue 155 as His
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match             67.3%;  Score 33;  DB 2;  Length 274;
  Best Local Similarity   87.5%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LTRTRGLK 10
              ||| ||||
Db        223 LTRKRGLK 230


RESULT 13
D70375
transcription termination factor Rho - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: D70375
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: D70375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-436 <AQF>
A;Cross-references: UNIPROT:O67031; GB:AE000711; NID:g2983401; PIDN:AAC06989.1; PID:g298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: rho
C;Superfamily: transcription termination factor rho
C;Keywords: transcription termination

  Query Match             67.3%;  Score 33;  DB 2;  Length 436;
  Best Local Similarity   87.5%;  Pred. No. 66;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LTRTRGLK 10
              |||| |||
Db         40 LTRTTGLK 47


RESULT 14
T49646
hsp70 related protein [imported] - Neurospora crassa
N;Alternate names: protein B5O22.280
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49646
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49646
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-449 <SCH>
A;Cross-references: UNIPROT:Q9P5S4; EMBL:AL355932; GSPDB:GN00116; NCSP:B5O22.280
A;Experimental source: BAC clone B5O22; strain OR74A
C;Genetics:
A;Gene: NCSP:B5O22.280
A;Map position: 6
A;Introns: 87/1; 161/2; 339/3

  Query Match             67.3%;  Score 33;  DB 2;  Length 449;
  Best Local Similarity   60.0%;  Pred. No. 68;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              |||| |:|::
Db        404 TRLTTTKGVE 413


RESULT 15
JH0101
apolipoprotein B-100 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0101; S33128; D60950
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Accession: JH0101
A;Molecule type: DNA
A;Residues: 1-784 <SMI>
A;Cross-references: UNIPROT:Q61314; GB:M35186
R;Smith, T.; Hautamaa, D.; Maeda, N.
submitted to the EMBL Data Library, May 1989
A;Reference number: S33128
A;Accession: S33128
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531,'S',533-784 <SM2>
A;Cross-references: EMBL:X15191
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LD
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: D60950
A;Molecule type: mRNA
A;Residues: 427-531,'S',533-700 <LAW>
C;Genetics:
A;Gene: MGI:Apob
A;Cross-references: MGI:88052
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein
```

F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

Query Match 67.3%; Score 33; DB 2; Length 784;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TRLTRTRGLK 10
              :|| | ||||
Db        647 SRLMRKRGLK 656
```

Search completed: December 29, 2004, 12:39:02
Job time : 10.6591 secs

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 58.4091 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-4
Perfect score: 49
Sequence: 1 TRLTRTRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	87.8	414	2	Q7YQR5	Q7yqr5 aotus vocif
2	43	87.8	596	2	Q28473	Q28473 macaca fasc
3	43	87.8	3262	2	Q13788	Q13788 homo sapien
4	43	87.8	4563	1	APB_HUMAN	P04114 homo sapien
5	43	87.8	4563	2	Q7Z600	Q7z600 homo sapien
6	39	79.6	421	2	Q7TN68	Q7tn68 glaucomys v
7	39	79.6	432	2	Q7YR10	Q7yr10 diceros bic
8	39	79.6	436	2	Q7YQM8	Q7yqm8 nyctimene a
9	39	79.6	438	2	Q7YQM7	Q7yqm7 pteropus hy
10	39	79.6	438	2	Q7YR04	Q7yr04 rousettus a
11	39	79.6	445	2	Q7YR08	Q7yr08 chaetophrac
12	39	79.6	445	2	Q7TN64	Q7tn64 agouti paca
13	39	79.6	445	2	Q7TN71	Q7tn71 hydrochoeru
14	39	79.6	445	2	Q7TN72	Q7tn72 erethizon d
15	39	79.6	780	2	Q60536	Q60536 mesocricetu
16	39	79.6	780	2	Q60537	Q60537 mesocricetu
17	38	77.6	141	2	Q8QUT5	Q8qut5 infectious
18	38	77.6	178	2	Q7TUB6	Q7tub6 prochloroco
19	37	75.5	275	2	Q7LZ77	Q7lz77 gallus gall
20	37	75.5	387	2	Q7YQN2	Q7yqn2 phalanger o
21	37	75.5	400	2	Q7YQM9	Q7yqm9 ornithorhyn
22	37	75.5	405	2	Q7YQN0	Q7yqn0 tachyglossu
23	37	75.5	445	2	Q7TN70	Q7tn70 dinomys bra
24	36	73.5	153	2	Q9FXM2	Q9fxm2 arabidopsis
25	36	73.5	202	2	Q8L8T0	Q8l8t0 arabidopsis
26	36	73.5	202	2	Q9LVA4	Q9lva4 arabidopsis
27	36	73.5	309	2	Q8XFV9	Q8xfv9 salmonella
28	36	73.5	309	2	Q7CPN5	Q7cpn5 salmonella
29	36	73.5	323	2	Q6TMM7	Q6tmm7 streptomyce
30	36	73.5	323	2	AAQ93596	Aaq93596 streptomy
31	36	73.5	407	2	Q7TN65	Q7tn65 atherurus a
32	36	73.5	412	2	Q7TN69	Q7tn69 hystrix bra
33	35	71.4	115	2	Q7ZET5	Q7zet5 human immun
34	35	71.4	165	2	Q8CCC7	Q8ccc7 mus musculu
35	35	71.4	289	2	Q8DUW3	Q8duw3 streptococc
36	35	71.4	317	2	Q6Z4N1	Q6z4n1 oryza sativ
37	35	71.4	317	2	BAC83809	Bac83809 oryza sat
38	35	71.4	522	2	Q88B51	Q88b51 pseudomonas
39	35	71.4	724	2	Q7Z405	Q7z405 homo sapien
40	35	71.4	724	2	Q7TN63	Q7tn63 mus musculu
41	35	71.4	851	2	Q6C4R4	Q6c4r4 yarrowia li
42	35	71.4	909	1	SBCC_DEIRA	Q9rt44 deinococcus
43	35	71.4	1059	2	Q7R4X6	Q7r4x6 giardia lam
44	35	71.4	1101	2	Q6N088	Q6n088 homo sapien
45	35	71.4	1101	2	CAE45782	Cae45782 homo sapi

ALIGNMENTS

RESULT 1
Q7YQR5
ID Q7YQR5 PRELIMINARY; PRT; 414 AA.
AC Q7YQR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein B 100 (Fragment).
GN Name=apoB-100;
OS Aotus vociferans (Spix's owl monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=57176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22761261; PubMed=12878460;
RA Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT "A new phylogenetic marker, apolipoprotein B, provides compelling
RT evidence for eutherian relationships.";
RL Mol. Phylogenet. Evol. 28:225-240(2003).
DR EMBL; AF548396; AAP97352.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45955 MW; EEFA8492157E1BDE CRC64;

Query Match 87.8%; Score 43; DB 2; Length 414;
Best Local Similarity 90.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRTRGLK 10
||||| ||||
Db 258 TRLTRKRGLK 267

RESULT 2
Q28473
ID Q28473 PRELIMINARY; PRT; 596 AA.
AC Q28473;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Apolipoprotein B (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92075708; PubMed=1742325;
RA Pape M.E., Castle C.K., Murray R.W., Funk G.M., Hunt C.E.,
RA Marotti K.R., Melchior G.W.;

```
RT   "Apo B metabolism in the cynomolgus monkey: evidence for post-
RT   transcriptional regulation.";
RL   Biochim. Biophys. Acta 1086:326-334(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Murray R.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X15737; CAA33755.1; -.
DR   PIR; S32802; S32802.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     596    596
SQ   SEQUENCE   596 AA;  66757 MW;  B13BBA74E25C3120 CRC64;

  Query Match             87.8%;  Score 43;  DB 2;  Length 596;
  Best Local Similarity   90.0%;  Pred. No. 2.7;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db        226 TRLTRKRGLK 235


RESULT 3
Q13788
ID   Q13788      PRELIMINARY;      PRT;  3262 AA.
AC   Q13788;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   APOB protein (Fragment).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191999; PubMed=2883086;
RA   Carlsson P., Darnfors C., Olofsson S.O., Bjursell G.;
RT   "Analysis of the human apolipoprotein B gene; complete structure of
RT   the B-74 region.";
RL   Gene 49:29-51(1986).
DR   EMBL; M15421; AAA51758.1; -.
DR   PIR; A27850; LPHUB.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005319; F:lipid transporter activity; NAS.
DR   GO; GO:0006869; P:lipid transport; NAS.
FT   NON_TER       1      1
SQ   SEQUENCE   3262 AA;  370140 MW;  56603BC0618DD40D CRC64;

  Query Match             87.8%;  Score 43;  DB 2;  Length 3262;
  Best Local Similarity   90.0%;  Pred. No. 18;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db       2084 TRLTRKRGLK 2093


RESULT 4
APB_HUMAN
ID   APB_HUMAN      STANDARD;      PRT;  4563 AA.
AC   P04114; O00502; Q13787;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein
DE   B-48 (Apo B-48)].
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87016385; PubMed=3763409;
RA   Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.,
RA   Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
RT   "Complete cDNA and derived protein sequence of human apolipoprotein B-
RT   100.";
RL   Nucleic Acids Res. 14:7501-7503(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=88003974; PubMed=3652907;
RA   Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
RA   Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
RT   "DNA sequence of the human apolipoprotein B gene.";
RL   DNA 6:363-372(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS ILE-98 AND GLU-4181.
RX   MEDLINE=87008488; PubMed=3759943;
RA   Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
RA   Gotto A.M. Jr., Chan L.;
RT   "The complete cDNA and amino acid sequence of human apolipoprotein B-
RT   100.";
RL   J. Biol. Chem. 261:12918-12921(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87041416; PubMed=3464946;
RA   Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
RA   Lee N., Brewer H.B. Jr.;
RT   "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
RT   derived amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87161758; PubMed=3030729;
RA   Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
RA   Zannis V.I.;
RT   "The complete sequence and structural analysis of human apolipoprotein
RT   B-100: relationship between apoB-100 and apoB-48 forms.";
RL   EMBO J. 5:3495-3507(1986).
RN   [6]
RP   SEQUENCE OF 709-906 FROM N.A.
RX   MEDLINE=85270450; PubMed=3860836;
RA   Deeb S.S., Motulsky A.G., Albers J.J.;
RT   "A partial cDNA clone for human apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
RN   [7]
RP   SEQUENCE OF 3056-3159 FROM N.A.
RX   MEDLINE=86041888; PubMed=3903660;
RA   Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
RA   Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
RT   "Human apolipoprotein B: identification of cDNA clones and
RT   characterization of mRNA.";
RL   Nucleic Acids Res. 13:6937-6953(1985).
RN   [8]
RP   SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=86093680; PubMed=3841204;
RA   Carlsson P., Olofsson S.O., Bondjers G., Darnfors C., Wiklund O.,
RA   Bjursell G.;
RT   "Molecular cloning of human apolipoprotein B cDNA.";
RL   Nucleic Acids Res. 13:8813-8826(1985).
RN   [9]
RP   SEQUENCE OF 3109-4563 FROM N.A.
RX   MEDLINE=85300528; PubMed=2994225;
RA   Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F., Urdea M.S.,
RA   Levy-Wilson B., Powell L.M., Pease R.J., Eddy R., Nakai H., Byers M.,
RA   Priestley L.M., Robertson E., Rall L.B., Betsholtz C., Shows T.B.,
RA   Mahley R.W., Scott J.;
RT   "Human apolipoprotein B: structure of carboxyl-terminal domains, sites
RT   of gene expression, and chromosomal localization.";
RL   Science 230:37-43(1985).
RN   [10]
```

```
RP   SEQUENCE OF 1-291 FROM N.A.
RX   MEDLINE=86149325; PubMed=3513177;
RA   Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
RA   Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
RT   "Isolation of a cDNA clone encoding the amino-terminal region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
RN   [11]
RP   SEQUENCE OF 1-1670 FROM N.A., AND VARIANT ILE-98.
RX   MEDLINE=86287319; PubMed=3461454;
RA   Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W., Yamanaka M.,
RA   Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
RT   "Analysis of cDNA clones encoding the entire B-26 region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
RN   [12]
RP   PARTIAL SEQUENCE, AND IDENTIFICATION OF APO-B48.
RX   MEDLINE=88018019; PubMed=3659919;
RA   Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
RA   Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
RA   Gotto A.M. Jr., Li W.-H., Chan L.;
RT   "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
RT   specific in-frame stop codon.";
RL   Science 238:363-366(1987).
RN   [13]
RP   DOMAINS.
RX   MEDLINE=87039351; PubMed=3773997;
RA   Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
RA   Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
RA   Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
RA   Levy-Wilson B., Scott J.;
RT   "Complete protein sequence and identification of structural domains of
RT   human apolipoprotein B.";
RL   Nature 323:734-738(1986).
RN   [14]
RP   DOMAINS.
RA   Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
RA   Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
RA   Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
RT   "Sequence, structure, receptor-binding domains and internal repeats of
RT   human apolipoprotein B-100.";
RL   Nature 323:738-742(1986).
RN   [15]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=86242245; PubMed=3087360;
RA   Dashti N., Lee D.M., Mok T.;
RT   "Apolipoprotein B is a calcium binding protein.";
RL   Biochem. Biophys. Res. Commun. 137:493-499(1986).
RN   [16]
RP   PALMITOYLATION OF CYS-1112.
RX   MEDLINE=20143590; PubMed=10679026;
RA   Zhao Y., McCabe J.B., Vance J., Berthiaume L.G.;
RT   "Palmitoylation of apolipoprotein B is required for proper
RT   intracellular sorting and transport of cholesteroyl esters and
RT   triglycerides.";
RL   Mol. Biol. Cell 11:721-734(2000).
RN   [17]
RP   VARIANT SER-4338.
RX   MEDLINE=91071750; PubMed=1979313;
RA   Navajas M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
RA   Cunny G., Cambien F., Roizes G.;
RT   "Detection by denaturing gradient gel electrophoresis of a new
RT   polymorphism in the apolipoprotein B gene.";
RL   Hum. Genet. 86:91-93(1990).
RN   [18]
RP   VARIANT FDB GLN-3527.
RX   MEDLINE=89098975; PubMed=2563166;
RA   Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
RA   McCarthy B.J.;
RT   "Association between a specific apolipoprotein B mutation and familial
RT   defective apolipoprotein B-100.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
RN   [19]
RP   VARIANT LEU-2739.
RX   MEDLINE=91016974; PubMed=2216805;
RA   Huang L.-S., Gavish D., Breslow J.L.;
RT   "Sequence polymorphism in the human apoB gene at position 8344.";
RL   Nucleic Acids Res. 18:5922-5922(1990).
RN   [20]
RP   VARIANT FDB CYS-3558.
RX   MEDLINE=95190020; PubMed=7883971;
RA   Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
RA   Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
RT   "Familial ligand-defective apolipoprotein B. Identification of a new
RT   mutation that decreases LDL receptor binding affinity.";
RL   J. Clin. Invest. 95:1225-1234(1995).
RN   [21]
RP   VARIANTS LEU-1437; SER-1914; LYS-2566; THR-3121; ALA-3945; MET-4128
RP   AND THR-4481.
RX   MEDLINE=97044521; PubMed=8889592;
RA   Poirier O., Ricard S., Behague I., Souriau C., Evans A.E.,
RA   Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
RT   "Detection of new variants in the apolipoprotein B (Apo B) gene by
RT   PCR-SSCP.";
RL   Hum. Mutat. 8:282-285(1996).
RN   [22]
RP   VARIANTS FDB GLN-3527 AND CYS-3558.
RX   MEDLINE=97403938; PubMed=9259199;
RA   Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
RA   Krempf M., Giraudet P., Junien C., Boileau C.;
RT   "Familial ligand-defective apolipoprotein B-100: simultaneous
RT   detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French
RT   population.";
RL   Hum. Mutat. 10:160-163(1997).
RN   [23]
RP   VARIANTS SER-1914; ARG-1923; LEU-2739; ASP-3319; THR-3427; GLN-3432
RP   AND ILE-3921.
RX   MEDLINE=98141125; PubMed=9490296;
RA   Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
RT   "Screening for mutations of the apolipoprotein B gene causing
RT   hypocholesterolemia.";
RL   Hum. Genet. 102:44-49(1998).
CC   -!- FUNCTION: Apolipoprotein B is a major protein constituent of
CC       chylomicrons (apo B-48), LDL (apo B-100) and VLDL (apo B-100). Apo
CC       B-100 functions as a recognition signal for the cellular binding
CC       and internalization of LDL particles by the apoB/E receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.

  Query Match             87.8%;  Score 43;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 26;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 5
Q7Z600
ID   Q7Z600      PRELIMINARY;      PRT;  4563 AA.
AC   Q7Z600;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Including Ag(X) antigen).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY324608; AAP72970.1; -.
DR   GO; GO:0005319; F:lipid transporter activity; IEA.
DR   GO; GO:0006869; P:lipid transport; IEA.
DR   InterPro; IPR009454; DUF1081.
DR   InterPro; IPR001747; Lipid_transprt_N.
DR   Pfam; PF06448; DUF1081; 1.
DR   Pfam; PF01347; Vitellogenin_N; 1.
DR   SMART; SM00638; LPD_N; 1.
KW   Lipoprotein.
SQ   SEQUENCE   4563 AA;  515553 MW;  030B34167CEDC63C CRC64;

  Query Match             87.8%;  Score 43;  DB 2;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 26;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 6
Q7TN68
ID   Q7TN68      PRELIMINARY;      PRT;   421 AA.
AC   Q7TN68;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Glaucomys volans (Southern flying squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Petauristinae;
OC   Glaucomys.
OX   NCBI_TaxID=64683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243379; AAP50767.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     421    421
SQ   SEQUENCE   421 AA;  46747 MW;  D47B77BD4F864FD1 CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 421;
  Best Local Similarity   80.0%;  Pred. No. 13;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        264 SRLTRKRGLK 273


RESULT 7
Q7YR10
ID   Q7YR10      PRELIMINARY;      PRT;   432 AA.
AC   Q7YR10;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243375; AAP50763.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     432    432
SQ   SEQUENCE   432 AA;  48171 MW;  F27B7AB39604732C CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 432;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        275 SRLTRKRGLK 284


RESULT 8
Q7YQM8
ID   Q7YQM8      PRELIMINARY;      PRT;   436 AA.
AC   Q7YQM8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Nyctimene albiventer (Common tube-nosed fruit bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Nyctimene.
OX   NCBI_TaxID=48988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548435; AAP97391.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     436    436
SQ   SEQUENCE   436 AA;  48717 MW;  1C4A7EAD72D2C629 CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 436;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        279 SRLTRKRGLK 288


RESULT 9
Q7YQM7
ID   Q7YQM7      PRELIMINARY;      PRT;   438 AA.
AC   Q7YQM7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Pteropus hypomelanus (Small flying fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Pteropus.
OX   NCBI_TaxID=9405;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
```

```
DR   EMBL; AF548436; AAP97392.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48734 MW;  2BD85BCBF4E2CC41 CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        281 SRLTRKRGLK 290


RESULT 10
Q7YR04
ID   Q7YR04      PRELIMINARY;      PRT;   438 AA.
AC   Q7YR04;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Rousettus amplexicaudatus (Common rousette).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Rousettus.
OX   NCBI_TaxID=58083;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243383; AAP50771.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48597 MW;  41C890DEAF95C872 CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        281 SRLTRKRGLK 290


RESULT 11
Q7YR08
ID   Q7YR08      PRELIMINARY;      PRT;   445 AA.
AC   Q7YR08;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Chaetophractus villosus (South American armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Chaetophractus.
OX   NCBI_TaxID=29080;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243378; AAP50766.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49564 MW;  2DA5DC3ED2F0FDD2 CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 12
Q7TN64
ID   Q7TN64      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN64;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Agouti paca (Paca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Agoutidae; Agouti.
OX   NCBI_TaxID=108852;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548417; AAP97373.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49721 MW;  34AF7ABE90F121EF CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 13
Q7TN71
ID   Q7TN71      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Hydrochoerus hydrochaeris (Capybara) (Carpincho).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Hydrochaeridae;
OC   Hydrochaeris.
OX   NCBI_TaxID=10149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243369; AAP50757.1; -.
DR   InterPro; IPR000871; Beta_lactamase_A.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49520 MW;  CBBA2DD53D7A18D2 CRC64;
```

```
  Query Match             79.6%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 14
Q7TN72
ID   Q7TN72      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN72;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Erethizon dorsatum (North American porcupine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Erethizontidae;
OC   Erethizon.
OX   NCBI_TaxID=34844;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243368; AAP50756.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49617 MW;  9572FE5F5E7625F2 CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 14;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 15
Q60536
ID   Q60536      PRELIMINARY;      PRT;   780 AA.
AC   Q60536;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hamster apolipoprotein (apoB) (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236327; PubMed=2332175;
RA   Smith T.J., Hautamaa D., Maeda N.;
RT   "Sequence of the putative low-density lipoprotein receptor-binding
RT   regions of apolipoprotein B in mouse and hamster.";
RL   Gene 87:309-310(1990).
DR   EMBL; M35187; AAA37059.1; -.
DR   PIR; C60950; C60950.
DR   PIR; JH0102; JH0102.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     780    780
SQ   SEQUENCE   780 AA;  86625 MW;  E371D1B2079D8F7E CRC64;

  Query Match             79.6%;  Score 39;  DB 2;  Length 780;
  Best Local Similarity   80.0%;  Pred. No. 27;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRTRGLK 10
              :|||| ||||
Db        642 SRLTRKRGLK 651


Search completed: December 29, 2004, 12:37:31
Job time : 59.5202 secs
```

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-5
Perfect score: 48
Sequence: 1 TRLTRSRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	10	2	AAY30686	Aay30686 Apo-B100
2	45	93.8	10	2	AAY30684	Aay30684 Apo-B100
3	45	93.8	10	2	AAY30685	Aay30685 Apo-B100
4	44	91.7	10	2	AAY30683	Aay30683 Apo-B100
5	44	91.7	10	2	AAY30682	Aay30682 Apo-B100
6	44	91.7	10	2	AAY30687	Aay30687 Apo-B100
7	44	91.7	11	2	AAW57205	Aaw57205 Apo B bin
8	44	91.7	13	2	AAW57207	Aaw57207 Apo B 100
9	44	91.7	15	2	AAW41261	Aaw41261 Apolipopr
10	44	91.7	15	2	AAW96892	Aaw96892 ApoB-100
11	44	91.7	20	6	ABJ37575	Abj37575 Heparin b
12	44	91.7	22	2	AAW57208	Aaw57208 Apo B 100
13	44	91.7	22	2	AAW57209	Aaw57209 Apo B 100
14	44	91.7	34	5	AAE14541	Aae14541 Human apo
15	44	91.7	36	2	AAW96876	Aaw96876 Nucleic a
16	44	91.7	37	2	AAW64587	Aaw64587 Human apo
17	44	91.7	51	2	AAW96845	Aaw96845 Nucleic a
18	44	91.7	343	4	ABB37687	Abb37687 Peptide #
19	44	91.7	343	4	ABG52504	Abg52504 Human liv
20	44	91.7	377	2	AAR72704	Aar72704 Human apo
21	44	91.7	377	2	AAR34031	Aar34031 Sequence
22	44	91.7	2463	8	ADJ57400	Adj57400 Human apo
23	44	91.7	3923	2	AAY31237	Aay31237 Human Apo
24	44	91.7	4536	2	AAW41262	Aaw41262 Apolipopr
25	44	91.7	4536	2	AAW96826	Aaw96826 Amino aci
26	44	91.7	4560	5	AAU98981	Aau98981 Human apo
27	44	91.7	4561	7	ADD48677	Add48677 Human Pro
28	44	91.7	4563	5	AAO15893	Aao15893 Human apo
29	44	91.7	4563	6	ABR40253	Abr40253 Human ali
30	44	91.7	4563	6	ABU79140	Abu79140 Apolipopr
31	44	91.7	4563	7	ADF43408	Adf43408 Apolipopr
32	44	91.7	4563	8	ADH18871	Adh18871 Human apo
33	44	91.7	4563	8	ADH18870	Adh18870 Human apo
34	44	91.7	4563	8	ADO33445	Ado33445 Human apo
35	44	91.7	4563	8	ADO33447	Ado33447 Human apo
36	44	91.7	4590	4	AAU33184	Aau33184 Novel hum
37	39	81.2	10	2	AAY30690	Aay30690 Apo-B100
38	39	81.2	10	2	AAY30692	Aay30692 Apo-B100
39	39	81.2	10	2	AAY30688	Aay30688 Apo-B100
40	39	81.2	11	2	AAW57206	Aaw57206 Apo B 100
41	39	81.2	11	2	AAW87717	Aaw87717 Analogue
42	39	81.2	11	5	AAE21732	Aae21732 BSMR effe
43	39	81.2	11	6	ABU07938	Abu07938 Apoprotei
44	39	81.2	11	7	ADF56451	Adf56451 Human apo
45	39	81.2	12	2	AAW41260	Aaw41260 Apolipopr

ALIGNMENTS

RESULT 1
AAY30686
ID AAY30686 standard; peptide; 10 AA.
XX
AC AAY30686;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 48; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||||||||||
Db 1 TRLTRSRGLK 10

RESULT 2
AAY30684
ID AAY30684 standard; peptide; 10 AA.
XX
AC AAY30684;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 93.8%; Score 45; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.029;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
|||||:||||
Db 1 TRLTRARGLK 10

RESULT 3
AAY30685
ID AAY30685 standard; peptide; 10 AA.
XX
AC AAY30685;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 93.8%; Score 45; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.029;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10

```
              |||||:||||
Db          1 TRLTRTRGLK 10


RESULT 4
AAY30683
ID   AAY30683 standard; peptide; 10 AA.
XX
AC   AAY30683;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             91.7%;  Score 44;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.047;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db          1 TRLTRDRGLK 10


RESULT 5
AAY30682
ID   AAY30682 standard; peptide; 10 AA.
XX
AC   AAY30682;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             91.7%;  Score 44;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.047;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db          1 TRLTRERGLK 10


RESULT 6
AAY30687
ID   AAY30687 standard; peptide; 10 AA.
XX
AC   AAY30687;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
```

KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 91.7%; Score 44; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.047;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||||| ||||
Db 1 TRLTRQRGLK 10

RESULT 7
AAW57205
ID AAW57205 standard; peptide; 11 AA.
XX
AC AAW57205;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B binding site peptide 2.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 12; Page 52; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B binding site
CC peptide which can be used as a component of a non-naturally occurring,
CC receptor-competent low density lipoprotein (LDL) particle of the present
CC invention. The LDL particle comprises at least 1 peptide component that
CC has at least 1 binding site for an apo B protein receptor and at least 1
CC lipophilic substituent. Also described in the invention are peptides
CC containing an apo B binding sequence with at least 70% identity with
CC sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC naturally occurring, receptor-competent LDL particles are useful as: (i)
CC drug-targeting vectors for delivering anticancer drugs to cancer cells
CC that express an apo B protein receptor, and (ii) additives for cell
CC culture media especially as growth supplements. Non-naturally occurring,
CC receptor-competent LDL particles do not require the complete apo B
CC sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 11 AA;

Query Match 91.7%; Score 44; DB 2; Length 11;
Best Local Similarity 90.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||||| ||||
Db 2 TRLTRKRGLK 11

RESULT 8
AAW57207
ID AAW57207 standard; peptide; 13 AA.
XX
AC AAW57207;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide B.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.

XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 13 AA;

Query Match 91.7%; Score 44; DB 2; Length 13;
Best Local Similarity 90.0%; Pred. No. 0.062;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||||| ||||
Db 3 TRLTRKRGLK 12

RESULT 9
AAW41261
ID AAW41261 standard; peptide; 15 AA.
XX
AC AAW41261;
XX
DT 19-MAY-1998 (first entry)
XX
DE Apolipoprotein B-100 fragment.
XX
KW Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW prothrombinase complex.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9743311-A1.
XX
PD 20-NOV-1997.
XX
PF 09-MAY-1997; 97WO-GB001255.
XX
PR 09-MAY-1996; 96GB-00009702.
XX
PA (UNLO) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI Bruckdorfer KR, Ettelaie C;
XX
DR WPI; 1998-008798/01.
XX
PT Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT used for treating or preventing coagulation, inhibiting angiogenesis,
PT cell differentiation and apoptosis.
XX
PS Disclosure; Page 22; 60pp; English.
XX
CC This sequence is an example of the peptide of the invention. It has the
CC formula (I), or their variants with one or more internal deletions,
CC insertions or substitutions, while retaining anti-coagulant properties of
CC apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC aa. Compositions containing the peptide are used for simultaneous,
CC separate or sequential treatment of cancer, particularly to prevent
CC metastatic spread. They are also used to inhibit thromboplastin-mediated
CC processes, specifically to prevent or reduce blood coagulation (e.g.
CC during or after surgery or in cases of heart attack, stroke etc.) and to
CC inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC which is active as such or as part of a 98-aa peptide, inhibits
CC activation of the prothrombinase complex; and prevents activation of
CC factor VII on the surface of thromboplastin and of platelets by thrombin.
CC It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC smaller than apoB-100, they act more quickly
XX
SQ Sequence 15 AA;

Query Match 91.7%; Score 44; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.071;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||||| ||||
Db 1 TRLTRKRGLK 10

RESULT 10
AAW96892
ID AAW96892 standard; peptide; 15 AA.
XX
AC AAW96892;
XX
DT 22-APR-1999 (first entry)
XX
DE ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from

```
CC   human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC   component of very-low density lipoproteins (VLDL), intermediate density
CC   lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC   present sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 15 AA;

  Query Match             91.7%;  Score 44;  DB 2;  Length 15;
  Best Local Similarity   90.0%;  Pred. No. 0.071;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db          6 TRLTRKRGLK 15


RESULT 11
ABJ37575
ID   ABJ37575 standard; peptide; 20 AA.
XX
AC   ABJ37575;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Heparin binding peptide sequence #28.
XX
KW   Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW   rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS   Unidentified.
XX
PN   WO2003007689-A2.
XX
PD   30-JAN-2003.
XX
PF   22-JUL-2002; 2002WO-US023419.
XX
PR   20-JUL-2001; 2001US-0306726P.
XX
PA   (ETHZ-) ETH ZUERICH.
PA   (UYZU-) UNIV ZURICH.
XX
PI   Hubbell JA,  Schoenmakers R,  Maynard HD;
XX
DR   WPI; 2003-300420/29.
XX
PT   Use of a ligand comprising of at least one sulfated or sulfonated amino
PT   acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT   retinopathy and hypoxia.
XX
PS   Disclosure; Fig 2; 79pp; English.
XX
CC   The invention relates to a novel ligand for binding a target biomolecule,
CC   which comprises a peptide having at least one sulphated or sulphonated
CC   amino acid and at least one amino acid chosen from neutral and positively
CC   charged amino acids. The novel ligands can be used for the treatment of
CC   e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC   This sequence represents a heparin binding peptide relating to the
CC   invention
XX
SQ   Sequence 20 AA;

  Query Match             91.7%;  Score 44;  DB 6;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 0.096;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 12
AAW57208
ID   AAW57208 standard; peptide; 22 AA.
XX
AC   AAW57208;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide C.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
FT   Modified-site   22
FT                   /note= "attached to cholesterol"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             91.7%;  Score 44;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16
```

```
RESULT 13
AAW57209
ID   AAW57209 standard; peptide; 22 AA.
XX
AC   AAW57209;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide D.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 22 AA;

  Query Match             91.7%;  Score 44;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db          7 TRLTRKRGLK 16


RESULT 14
AAE14541
ID   AAE14541 standard; peptide; 34 AA.
XX
AC   AAE14541;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human apoB-100 derived peptide p62.
XX
KW   Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW   cardiovascular disease; coronary heart disease; pre-eclampsia;
KW   non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW   peptide p62.
XX
OS   Homo sapiens.
XX
PN   WO200206314-A2.
XX
PD   24-JAN-2002.
XX
PF   18-JUL-2001; 2001WO-GB003212.
XX
PR   18-JUL-2000; 2000GB-00017641.
XX
PA   (ARKT-) ARK THERAPEUTICS LTD.
XX
PI   Narvanen O,  Yla-Herttuala S;
XX
DR   WPI; 2002-179777/23.
XX
PT   New peptide useful in enzyme immunoassays for detecting oxidized low
PT   density lipoprotein which is a marker of coronary heart disease and other
PT   cardiovascular diseases, has affinity for oxidized low density
PT   lipoprotein.
XX
PS   Claim 6; Page 5; 21pp; English.
XX
CC   The invention relates to peptides having affinity for oxidised low
CC   density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC   is useful in an immunoassay to determine the presence, and optionally,
CC   the amount of antibodies in a sample, having affinity for oxLDL.
CC   Preferably immobilised peptide is useful for measuring the amount of
CC   autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC   from a patient for evaluating the risk of coronary heart diseases, other
CC   cardiovascular diseases, and several other disorders such as
CC   periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC   endothelial dysfunction. The peptide of the invention is stable, can be
CC   synthesised easily without the need to isolate proteins from a patient's
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             91.7%;  Score 44;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.17;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db         25 TRLTRKRGLK 34


RESULT 15
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
```

XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 16; Fig 12C; 293pp; English.
XX
CC AAW96827-77 represent nucleic acid binding domains derived from human
CC apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 36 AA;

Query Match 91.7%; Score 44; DB 2; Length 36;
Best Local Similarity 90.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db         11 TRLTRKRGLK 20
```

Search completed: December 29, 2004, 12:28:48
Job time : 61.0227 secs

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:28:58 ; Search time 8.97727 Seconds
(without alignments)
51.960 Million cell updates/sec

Title: US-09-823-418-5
Perfect score: 48
Sequence: 1 TRLTRSRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 187395 seqs, 46645940 residues

Total number of hits satisfying chosen parameters: 187395

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	91.7	4560	6	US-10-398-200-2	Sequence 2, Appli
2	44	91.7	4563	6	US-10-868-577A-25	Sequence 25, Appl
3	38	79.2	289	1	PCT-US02-09107B-72046	Sequence 72046, A
4	33	68.8	62	7	US-11-001-793-5628	Sequence 5628, Ap
5	33	68.8	470	7	US-11-001-793-5939	Sequence 5939, Ap
6	32	66.7	309	6	US-10-482-526A-576	Sequence 576, App
7	32	66.7	388	1	PCT-US04-35137-70	Sequence 70, Appl
8	32	66.7	388	6	US-10-972-963-70	Sequence 70, Appl
9	32	66.7	441	6	US-10-990-328-8115	Sequence 8115, Ap
10	32	66.7	441	6	US-10-990-328-8116	Sequence 8116, Ap
11	31	64.6	474	6	US-10-511-989-167	Sequence 167, App
12	31	64.6	522	1	PCT-US02-09107B-69335	Sequence 69335, A
13	31	64.6	533	1	PCT-US02-09107B-52732	Sequence 52732, A
14	31	64.6	534	1	PCT-US02-09107B-74744	Sequence 74744, A
15	31	64.6	535	1	PCT-US02-09107B-51874	Sequence 51874, A
16	31	64.6	535	1	PCT-US02-09107B-73828	Sequence 73828, A
17	31	64.6	536	1	PCT-US02-09107B-42481	Sequence 42481, A
18	31	64.6	536	1	PCT-US02-09107B-57662	Sequence 57662, A
19	31	64.6	536	1	PCT-US02-09107B-72504	Sequence 72504, A
20	31	64.6	657	6	US-10-990-328-13595	Sequence 13595, A
21	31	64.6	712	6	US-10-990-328-13596	Sequence 13596, A
22	31	64.6	720	6	US-10-990-328-11685	Sequence 11685, A
23	31	64.6	720	6	US-10-990-328-11686	Sequence 11686, A
24	31	64.6	720	6	US-10-990-328-11687	Sequence 11687, A
25	31	64.6	720	6	US-10-990-328-11688	Sequence 11688, A
26	31	64.6	933	1	PCT-US02-09107B-66243	Sequence 66243, A
27	31	64.6	935	6	US-10-511-989-14	Sequence 14, Appl
28	31	64.6	1012	6	US-10-511-989-12	Sequence 12, Appl
29	31	64.6	6995	1	PCT-US04-38392-4	Sequence 4, Appli
30	31	64.6	6995	6	US-10-983-340-4	Sequence 4, Appli
31	30	62.5	80	6	US-10-220-366A-22245	Sequence 22245, A
32	30	62.5	144	6	US-10-959-539-30	Sequence 30, Appl
33	30	62.5	148	6	US-10-408-765-840	Sequence 840, App
34	30	62.5	154	6	US-10-399-103A-561	Sequence 561, App
35	30	62.5	154	6	US-10-408-765-1915	Sequence 1915, Ap
36	30	62.5	154	6	US-10-408-765-2288	Sequence 2288, Ap
37	30	62.5	160	1	PCT-US02-09107B-60931	Sequence 60931, A
38	30	62.5	238	1	PCT-US04-37204-5640	Sequence 5640, Ap
39	30	62.5	286	1	PCT-US02-09107B-45389	Sequence 45389, A
40	30	62.5	296	1	PCT-US02-09107B-50307	Sequence 50307, A
41	30	62.5	297	6	US-10-408-765-2841	Sequence 2841, Ap
42	30	62.5	299	1	PCT-US02-09107B-45979	Sequence 45979, A
43	30	62.5	306	1	PCT-US02-09107B-77577	Sequence 77577, A
44	30	62.5	330	1	PCT-US02-09107B-74223	Sequence 74223, A
45	30	62.5	407	6	US-10-732-923-9577	Sequence 9577, Ap

ALIGNMENTS

RESULT 1
US-10-398-200-2
; Sequence 2, Application US/10398200
; GENERAL INFORMATION:
; APPLICANT: AGNELLO, VINCENT
; TITLE OF INVENTION: METHOD OF INHIBITING INFECTION BY HCV, OTHER
; TITLE OF INVENTION: FLAVIVIRIDAE VIRUSES, AND ANY OTHER VIRUS THAT
; TITLE OF INVENTION: COMPLEXES TO LOW DENSITY LIPOPROTEIN OR TO VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPRTOEIN IN BLOOD BY PREVENTING VIRAL ENTRY
; TITLE OF INVENTION: INTO A CELL
; FILE REFERENCE: 1513-PCT-00
; CURRENT APPLICATION NUMBER: US/10/398,200
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/243,594
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-398-200-2

Query Match 91.7%; Score 44; DB 6; Length 4560;
Best Local Similarity 90.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||||| ||||
Db 3382 TRLTRKRGLK 3391

RESULT 2
US-10-868-577A-25
; Sequence 25, Application US/10868577A
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al.
; TITLE OF INVENTION: HEPARIN BINDING VEGFR-3 LIGANDS
; FILE REFERENCE: 28967/39359A
; CURRENT APPLICATION NUMBER: US/10/868,577A
; CURRENT FILING DATE: 2004-06-14
; PRIOR APPLICATION NUMBER: US 60/478,390
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 10/669,176
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25

```
;   LENGTH: 4563
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(126)
;   OTHER INFORMATION: heparin binding domain
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3161)..(3236)
;   OTHER INFORMATION: heparin binding domain
US-10-868-577A-25

  Query Match             91.7%;  Score 44;  DB 6;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 3;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 3
PCT-US02-09107B-72046
; Sequence 72046, Application PC/TUS0209107B
; GENERAL INFORMATION:
;  APPLICANT: Elitra Pharmaceuticals Inc.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034VPC
;  CURRENT APPLICATION NUMBER: PCT/US02/09107B
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/815,242
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 09/948,993
;  PRIOR FILING DATE: 2001-09-06
;  PRIOR APPLICATION NUMBER: 60/342,923
;  PRIOR FILING DATE: 2001-10-25
;  PRIOR APPLICATION NUMBER: 10/072,851
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/362,699
;  PRIOR FILING DATE: 2002-03-06
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72046
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Streptococcus mutans
PCT-US02-09107B-72046

  Query Match             79.2%;  Score 38;  DB 1;  Length 289;
  Best Local Similarity   88.9%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTRSRGLK 10
              || ||||||
Db        241 RLVRSRGLK 249


RESULT 4
US-11-001-793-5628
; Sequence 5628, Application US/11001793
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et al.
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS900
;  CURRENT APPLICATION NUMBER: US/11/001,793
;  CURRENT FILING DATE:  2004-12-02
;  PRIOR APPLICATION NUMBER: US/10/100,683
;  PRIOR FILING DATE: 2002-03-19
;  PRIOR APPLICATION NUMBER: US 60/040,162
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: US 60/043,576
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: US 60/047,601
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: US 60/056,845
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/043,580
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: US 60/047,599
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: US 60/056,664
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/043,314
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: US 60/047,632
;  PRIOR FILING DATE: 1997-05-23
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13468
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5628
;   LENGTH: 62
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-001-793-5628

  Query Match             68.8%;  Score 33;  DB 7;  Length 62;
  Best Local Similarity   87.5%;  Pred. No. 5.9;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRG 8
              ||| ||||
Db         24 TRLRRSRG 31


RESULT 5
US-11-001-793-5939
; Sequence 5939, Application US/11001793
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et al.
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS900
;  CURRENT APPLICATION NUMBER: US/11/001,793
;  CURRENT FILING DATE:  2004-12-02
;  PRIOR APPLICATION NUMBER: US/10/100,683
;  PRIOR FILING DATE: 2002-03-19
;  PRIOR APPLICATION NUMBER: US 60/040,162
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: US 60/043,576
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: US 60/047,601
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: US 60/056,845
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/043,580
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: US 60/047,599
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: US 60/056,664
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/043,314
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: US 60/047,632
;  PRIOR FILING DATE: 1997-05-23
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13468
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5939
;   LENGTH: 470
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-001-793-5939

  Query Match             68.8%;  Score 33;  DB 7;  Length 470;
  Best Local Similarity   87.5%;  Pred. No. 48;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 TRLTRSRG 8
              ||| ||||
Db        432 TRLRRSRG 439


RESULT 6
US-10-482-526A-576
; Sequence 576, Application US/10482526A
; GENERAL INFORMATION:
;  APPLICANT: Syngenta Participations AG
;  TITLE OF INVENTION: PLANT DISEASE RESISTANCE GENES
;  FILE REFERENCE: S-70034A
;  CURRENT APPLICATION NUMBER: US/10/482,526A
;  CURRENT FILING DATE:  2003-12-18
;  PRIOR APPLICATION NUMBER: US 60/300,112
;  PRIOR FILING DATE: 2001-06-22
;  PRIOR APPLICATION NUMBER: US 60/325,277
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: US 60/366,535
;  PRIOR FILING DATE: 2002-03-22
;  NUMBER OF SEQ ID NOS: 1394
;  SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 576
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-10-482-526A-576

  Query Match             66.7%;  Score 32;  DB 6;  Length 309;
  Best Local Similarity   77.8%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGL 9
              ||  |||||
Db         43 TRTPRSRGL 51


RESULT 7
PCT-US04-35137-70
; Sequence 70, Application PC/TUS0435137
; GENERAL INFORMATION:
;  APPLICANT: Gencia Corporation
;  APPLICANT:  Khan, Shaharyar
;  TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides
;  FILE REFERENCE: 120701-2030
;  CURRENT APPLICATION NUMBER: PCT/US04/35137
;  CURRENT FILING DATE:  2004-11-03
;  PRIOR APPLICATION NUMBER: 60/568,436
;  PRIOR FILING DATE: 2004-05-05
;  PRIOR APPLICATION NUMBER: 60/513,983
;  PRIOR FILING DATE: 2003-10-24
;  NUMBER OF SEQ ID NOS: 218
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70
;   LENGTH: 388
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US04-35137-70

  Query Match             66.7%;  Score 32;  DB 1;  Length 388;
  Best Local Similarity   70.0%;  Pred. No. 63;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              | |||: |||
Db         52 TVLTRAEGLK 61


RESULT 8
US-10-972-963-70
; Sequence 70, Application US/10972963
; GENERAL INFORMATION:
;  APPLICANT: Gencia Corporation
;  APPLICANT:  Khan, Shaharyar
;  TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides
;  FILE REFERENCE: 120701-2030
;  CURRENT APPLICATION NUMBER: US/10/972,963
;  CURRENT FILING DATE:  2004-10-25
;  PRIOR APPLICATION NUMBER: 60/568,436
;  PRIOR FILING DATE: 2004-05-05
;  PRIOR APPLICATION NUMBER: 60/513,983
;  PRIOR FILING DATE: 2003-10-24
;  NUMBER OF SEQ ID NOS: 218
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70
;   LENGTH: 388
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-972-963-70

  Query Match             66.7%;  Score 32;  DB 6;  Length 388;
  Best Local Similarity   70.0%;  Pred. No. 63;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              | |||: |||
Db         52 TVLTRAEGLK 61


RESULT 9
US-10-990-328-8115
; Sequence 8115, Application US/10990328
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
;  TITLE OF INVENTION:  ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: CL001495
;  CURRENT APPLICATION NUMBER: US/10/990,328
;  CURRENT FILING DATE:  2004-11-17
;  NUMBER OF SEQ ID NOS: 558824
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8115
;   LENGTH: 441
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-990-328-8115

  Query Match             66.7%;  Score 32;  DB 6;  Length 441;
  Best Local Similarity   70.0%;  Pred. No. 72;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              | |||: |||
Db        105 TVLTRAEGLK 114


RESULT 10
US-10-990-328-8116
; Sequence 8116, Application US/10990328
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
;  TITLE OF INVENTION:  ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: CL001495
;  CURRENT APPLICATION NUMBER: US/10/990,328
;  CURRENT FILING DATE:  2004-11-17
;  NUMBER OF SEQ ID NOS: 558824
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8116
;   LENGTH: 441
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-990-328-8116
```

```
  Query Match             66.7%;  Score 32;  DB 6;  Length 441;
  Best Local Similarity   70.0%;  Pred. No. 72;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              | |||: |||
Db        105 TVLTRAEGLK 114


RESULT 11
US-10-511-989-167
; Sequence 167, Application US/10511989
; GENERAL INFORMATION:
;  APPLICANT: University of North Carolina-Chapel Hill
;  APPLICANT:  Ting, Jenny
;  APPLICANT:  Linhoff, Michael
;  APPLICANT:  Harton, Johnathan
;  APPLICANT:  Williams, Kristi
;  APPLICANT:  Lich, John
;  APPLICANT:  O'Connor, William
;  APPLICANT:  Moore, Christopher
;  APPLICANT:  Davis, Beckley
;  APPLICANT:  Brickey, W. Jane
;  APPLICANT:  Conti, Brian
;  APPLICANT:  Zhang, Jinghua
;  APPLICANT:  Zhu, Xin-Sheng
;  TITLE OF INVENTION: CATERPILLER GENE FAMILY
;  FILE REFERENCE: 5470.368WO
;  CURRENT APPLICATION NUMBER: US/10/511,989
;  CURRENT FILING DATE:  2004-10-21
;  PRIOR APPLICATION NUMBER: US 60/376,626
;  PRIOR FILING DATE: 2002-04-30
;  NUMBER OF SEQ ID NOS: 186
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 167
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-511-989-167

  Query Match             64.6%;  Score 31;  DB 6;  Length 474;
  Best Local Similarity   70.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              |||| |: ||
Db        146 TRLTTSKRLK 155


RESULT 12
PCT-US02-09107B-69335
; Sequence 69335, Application PC/TUS0209107B
; GENERAL INFORMATION:
;  APPLICANT: Elitra Pharmaceuticals Inc.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034VPC
;  CURRENT APPLICATION NUMBER: PCT/US02/09107B
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/815,242
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 09/948,993
;  PRIOR FILING DATE: 2001-09-06
;  PRIOR APPLICATION NUMBER: 60/342,923
;  PRIOR FILING DATE: 2001-10-25
;  PRIOR APPLICATION NUMBER: 10/072,851
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/362,699
;  PRIOR FILING DATE: 2002-03-06
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69335
;   LENGTH: 522
;   TYPE: PRT
;   ORGANISM: Pseudomonas syringae
PCT-US02-09107B-69335

  Query Match             64.6%;  Score 31;  DB 1;  Length 522;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||| |: ||:
Db        452 TRLQRTEGLR 461


RESULT 13
PCT-US02-09107B-52732
; Sequence 52732, Application PC/TUS0209107B
; GENERAL INFORMATION:
;  APPLICANT: Elitra Pharmaceuticals Inc.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034VPC
;  CURRENT APPLICATION NUMBER: PCT/US02/09107B
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/815,242
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 09/948,993
;  PRIOR FILING DATE: 2001-09-06
;  PRIOR APPLICATION NUMBER: 60/342,923
;  PRIOR FILING DATE: 2001-10-25
;  PRIOR APPLICATION NUMBER: 10/072,851
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/362,699
;  PRIOR FILING DATE: 2002-03-06
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52732
;   LENGTH: 533
;   TYPE: PRT
;   ORGANISM: Clostridium botulinum
PCT-US02-09107B-52732

  Query Match             64.6%;  Score 31;  DB 1;  Length 533;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTRSRGLK 10
              || ::||||
Db         27 RLLKNRGLK 35


RESULT 14
PCT-US02-09107B-74744
; Sequence 74744, Application PC/TUS0209107B
; GENERAL INFORMATION:
;  APPLICANT: Elitra Pharmaceuticals Inc.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034VPC
;  CURRENT APPLICATION NUMBER: PCT/US02/09107B
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/815,242
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 09/948,993
;  PRIOR FILING DATE: 2001-09-06
;  PRIOR APPLICATION NUMBER: 60/342,923
;  PRIOR FILING DATE: 2001-10-25
;  PRIOR APPLICATION NUMBER: 10/072,851
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/362,699
;  PRIOR FILING DATE: 2002-03-06
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74744
;   LENGTH: 534
;   TYPE: PRT
```

```
;   ORGANISM: Streptococcus pyogenes
PCT-US02-09107B-74744

  Query Match             64.6%;  Score 31;  DB 1;  Length 534;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTRSRGLK 10
              || ::||||
Db         26 RLLKNRGLK 34


RESULT 15
PCT-US02-09107B-51874
; Sequence 51874, Application PC/TUS0209107B
; GENERAL INFORMATION:
;  APPLICANT: Elitra Pharmaceuticals Inc.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034VPC
;  CURRENT APPLICATION NUMBER: PCT/US02/09107B
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/815,242
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 09/948,993
;  PRIOR FILING DATE: 2001-09-06
;  PRIOR APPLICATION NUMBER: 60/342,923
;  PRIOR FILING DATE: 2001-10-25
;  PRIOR APPLICATION NUMBER: 10/072,851
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/362,699
;  PRIOR FILING DATE: 2002-03-06
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51874
;   LENGTH: 535
;   TYPE: PRT
;   ORGANISM: Clostridium acetobutylicum
PCT-US02-09107B-51874

  Query Match             64.6%;  Score 31;  DB 1;  Length 535;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTRSRGLK 10
              || ::||||
Db         27 RLLKNRGLK 35


Search completed: December 29, 2004, 13:18:14
Job time : 9.97727 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 9.65909 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-5
Perfect score: 48
Sequence: 1 TRLTRSRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	91.7	596	2	S32802	apolipoprotein B -
2	44	91.7	4563	1	LPHUB	apolipoprotein B-1
3	40	83.3	269	2	C60950	apolipoprotein B-1
4	40	83.3	779	2	JH0102	apolipoprotein B -
5	38	79.2	275	2	E60950	apolipoprotein B-1
6	37	77.1	309	2	AH0906	conserved hypothet
7	35	72.9	290	2	S39854	traX protein - Str
8	35	72.9	309	1	E65112	hypothetical 34.6
9	35	72.9	309	2	E85985	hypothetical prote
10	35	72.9	309	2	B91140	hypothetical prote
11	34	70.8	173	2	G87383	acetyltransferase,
12	34	70.8	274	2	A60950	apolipoprotein B-1
13	34	70.8	614	1	S75294	ferrous iron trans
14	34	70.8	784	2	JH0101	apolipoprotein B-1
15	33	68.8	101	2	E72691	hypothetical prote
16	33	68.8	393	2	S48288	probable phosphopr
17	33	68.8	484	2	S40051	starch synthase (E
18	33	68.8	1838	2	T18448	pathogenicity fact
19	32	66.7	208	2	E72514	hypothetical prote
20	32	66.7	232	1	S28609	phosphoadenylyl-su
21	32	66.7	304	2	A98146	probable threonin
22	32	66.7	336	2	AC3142	threonine dehydrat
23	32	66.7	388	1	DEHUPT	pyruvate dehydroge
24	32	66.7	412	2	E83061	hypothetical prote
25	32	66.7	420	2	B72386	hypothetical prote
26	32	66.7	486	2	T40901	ranbpm homolog - f
27	32	66.7	487	1	LQBP34	DNA ligase (ATP) (
28	32	66.7	487	2	S06464	DNA ligase (ATP) (
29	32	66.7	506	2	AD3338	cobyric acid synth
30	32	66.7	680	2	AB1875	hypothetical prote
31	32	66.7	1036	2	S76027	hypothetical prote
32	31	64.6	146	2	T14681	myc-like regulator
33	31	64.6	233	1	C48560	UL56 protein - hum
34	31	64.6	272	2	E83363	hypothetical prote
35	31	64.6	290	2	E84797	hypothetical prote
36	31	64.6	329	2	T17033	leucine rich repea
37	31	64.6	330	1	F69534	pyruvate formate-l
38	31	64.6	398	2	D96795	probable DnaJ prot
39	31	64.6	535	2	C95057	CTP synthase [impo
40	31	64.6	535	2	C86686	CTP synthetase [im
41	31	64.6	535	2	G97255	CTP synthase (UTP-
42	31	64.6	535	2	F97926	CTP synthase (EC 6
43	31	64.6	712	2	S71626	3',5'-cyclic-nucle
44	31	64.6	732	2	T05448	hypothetical prote
45	31	64.6	788	2	S67595	hypothetical prote

ALIGNMENTS

RESULT 1
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchi
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional r
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g9301
C;Superfamily: apolipoprotein B

Query Match 91.7%; Score 44; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db        226 TRLTRKRGLK 235
```

RESULT 2
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A2
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Sc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNI
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: r
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986

A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732

R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.,
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than on
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place c
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

  Query Match 91.7%; Score 44; DB 1; Length 4563;
  Best Local Similarity 90.0%; Pred. No. 3.4;
  Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
 |||||.||||
Db 3385 TRLTRKRGLK 3394

RESULT 3
C60950
apolipoprotein B-100 - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: C60950
A;Molecule type: DNA
A;Residues: 1-269 <LAW>
A;Cross-references: UNIPROT:Q60537; UNIPROT:Q60536
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match 83.3%; Score 40; DB 2; Length 269;
  Best Local Similarity 80.0%; Pred. No. 1.5;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
 :|||| ||||
Db 216 SRLTRKRGLK 225

RESULT 4
JH0102
apolipoprotein B - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0102
R;Smith, T.J.
submitted to GenBank, June 1990
A;Reference number: A38864
A;Accession: JH0102
A;Molecule type: DNA
A;Residues: 1-779 <SMI>
A;Cross-references: UNIPROT:Q60536; GB:M35187
A;Note: this is a revision to the sequence from reference JH0101
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Contents: annotation
A;Note: this sequence has been revised in reference A38864
C;Genetics:
A;Gene: apoB
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

  Query Match 83.3%; Score 40; DB 2; Length 779;
  Best Local Similarity 80.0%; Pred. No. 4.1;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
 :|||| ||||
Db 642 SRLTRKRGLK 651

RESULT 5
E60950
apolipoprotein B-100 - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: E60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LD
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: E60950
A;Molecule type: mRNA
A;Residues: 1-275 <LAW>
A;Cross-references: UNIPROT:Q7LZ77
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein

  Query Match 79.2%; Score 38; DB 2; Length 275;
  Best Local Similarity 80.0%; Pred. No. 3.9;
  Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
 | ||| ||||
Db 221 TSLTRKRGLK 230

RESULT 6
AH0906
conserved hypothetical protein STY3508 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0906
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

```
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0906
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07846.1; PID:g16504394; GSPDB:GN00176
C;Genetics:
A;Gene: STY3508
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match             77.1%;  Score 37;  DB 2;  Length 309;
  Best Local Similarity   70.0%;  Pred. No. 7;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||: | ||||
Db        170 TRIARERGLK 179


RESULT 7
S39854
traX protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S39854; S32232
R;Brolle, D.F.; Pape, H.; Hopwood, D.A.; Kieser, T.
Mol. Microbiol. 10, 157-170, 1993
A;Title: Analysis of the transfer region of the Streptomyces plasmid SCP2*.
A;Reference number: S39853; MUID:95058174; PMID:7968512
A;Accession: S39854
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <BRO>
A;Cross-references: UNIPROT:Q06258; EMBL:X72857; NID:g288432; PIDN:CAA51379.1; PID:g5816
C;Genetics:
A;Start codon: GTG

  Query Match             72.9%;  Score 35;  DB 2;  Length 290;
  Best Local Similarity   87.5%;  Pred. No. 17;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTRSRGLK 10
              |||||||:
Db        130 LTRSRGLR 137


RESULT 8
E65112
hypothetical 34.6 kD protein in arcB-gltB intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E65112
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E65112
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-309 <BLAT>
A;Cross-references: UNIPROT:P45476; GB:AE000400; GB:U00096; NID:g2367203; PIDN:AAC76243.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match             72.9%;  Score 35;  DB 1;  Length 309;
  Best Local Similarity   70.0%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              |:| | ||||
Db        170 TQLARQRGLK 179


RESULT 9
E85985
hypothetical protein yhcC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85985
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85985
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <STO>
A;Cross-references: UNIPROT:P45476; GB:AE005174; NID:g12517832; PIDN:AAG58345.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match             72.9%;  Score 35;  DB 2;  Length 309;
  Best Local Similarity   70.0%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              |:| | ||||
Db        170 TQLARQRGLK 179


RESULT 10
B91140
hypothetical protein ECs4090 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B91140
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91140
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <HAY>
A;Cross-references: UNIPROT:P45476; GB:BA000007; PIDN:BAB37513.1; PID:g13363563; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4090
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match             72.9%;  Score 35;  DB 2;  Length 309;
  Best Local Similarity   70.0%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              |:| | ||||
Db        170 TQLARQRGLK 179


RESULT 11
G87383
acetyltransferase, GNAT family [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87383
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
```

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87383
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <STO>
A;Cross-references: UNIPROT:Q9A9B1; GB:AE005673; NID:g13422385; PIDN:AAK23067.1; GSPDB:G
C;Genetics:
A;Gene: CC1083

Query Match 70.8%; Score 34; DB 2; Length 173;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGL 9
||| |:|||
Db 49 TRLMRARGL 57

RESULT 12
A60950
apolipoprotein B-100 - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: A60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: A60950
A;Molecule type: mRNA
A;Residues: 1-274 <LAW>
A;Cross-references: UNIPROT:Q7M2U9
A;Note: authors translated the codon GAT for residue 155 as His
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

Query Match 70.8%; Score 34; DB 2; Length 274;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LTRSRGLK 10
||| ||||
Db 223 LTRKRGLK 230

RESULT 13
S75294
ferrous iron transport protein B - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein slr1392
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S75294
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75294
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-614 <KAN>
A;Cross-references: UNIPROT:P73182; EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BAA1720
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: feoB
C;Superfamily: ferrous iron transport protein B; translation elongation factor Tu homolo
C;Keywords: GTP binding; nucleotide binding; P-loop
F;19-134/Domain: translation elongation factor Tu homology <ETU>
F;25-32/Region: nucleotide-binding motif A (P-loop)

Query Match 70.8%; Score 34; DB 1; Length 614;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
||: ||||::
Db 361 TRVMRSRGMR 370

RESULT 14
JH0101
apolipoprotein B-100 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0101; S33128; D60950
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Accession: JH0101
A;Molecule type: DNA
A;Residues: 1-784 <SMI>
A;Cross-references: UNIPROT:Q61314; GB:M35186
R;Smith, T.; Hautamaa, D.; Maeda, N.
submitted to the EMBL Data Library, May 1989
A;Reference number: S33128
A;Accession: S33128
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531,'S',533-784 <SM2>
A;Cross-references: EMBL:X15191
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: D60950
A;Molecule type: mRNA
A;Residues: 427-531,'S',533-700 <LAW>
C;Genetics:
A;Gene: MGI:Apob
A;Cross-references: MGI:88052
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein,
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

Query Match 70.8%; Score 34; DB 2; Length 784;
Best Local Similarity 70.0%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRSRGLK 10
:|| | ||||
Db 647 SRLMRKRGLK 656

RESULT 15
E72691
hypothetical protein APE0949 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E72691
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takal
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72691
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-101 <KAW>
A;Cross-references: UNIPROT:Q9YDG4; DDBJ:AP000060; NID:g5104188; PIDN:BAA79933.1; PID:g
A;Experimental source: strain K1

C;Genetics:
A;Gene: APE0949
C;Superfamily: Aeropyrum pernix hypothetical protein APE0949

Query Match 68.8%; Score 33; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 LTRSRGL 9
              |||||||
Db         74 LTRSRGL 80
```

Search completed: December 29, 2004, 12:39:03
Job time : 10.6591 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 58.4091 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-5
Perfect score: 48
Sequence: 1 TRLTRSRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	91.7	414	2	Q7YQR5	Q7yqr5 aotus vocif
2	44	91.7	596	2	Q28473	Q28473 macaca fasc
3	44	91.7	3262	2	Q13788	Q13788 homo sapien
4	44	91.7	4563	1	APB_HUMAN	P04114 homo sapien
5	44	91.7	4563	2	Q7Z600	Q7z600 homo sapien
6	40	83.3	421	2	Q7TN68	Q7tn68 glaucomys v
7	40	83.3	432	2	Q7YR10	Q7yr10 diceros bic
8	40	83.3	436	2	Q7YQM8	Q7yqm8 nyctimene a
9	40	83.3	438	2	Q7YQM7	Q7yqm7 pteropus hy
10	40	83.3	438	2	Q7YR04	Q7yr04 rousettus a
11	40	83.3	445	2	Q7YR08	Q7yr08 chaetophrac
12	40	83.3	445	2	Q7TN64	Q7tn64 agouti paca
13	40	83.3	445	2	Q7TN71	Q7tn71 hydrochoeru
14	40	83.3	445	2	Q7TN72	Q7tn72 erethizon d
15	40	83.3	780	2	Q60536	Q60536 mesocricetu
16	40	83.3	780	2	Q60537	Q60537 mesocricetu
17	38	79.2	275	2	Q7LZ77	Q7lz77 gallus gall
18	38	79.2	289	2	Q8DUW3	Q8duw3 streptococc
19	38	79.2	387	2	Q7YQN2	Q7yqn2 phalanger o
20	38	79.2	400	2	Q7YQM9	Q7yqm9 ornithorhyn
21	38	79.2	405	2	Q7YQN0	Q7yqn0 tachyglossu
22	38	79.2	445	2	Q7TN70	Q7tn70 dinomys bra
23	37	77.1	309	2	Q8XFV9	Q8xfv9 salmonella
24	37	77.1	309	2	Q7CPN5	Q7cpn5 salmonella
25	37	77.1	407	2	Q7TN65	Q7tn65 atherurus a
26	37	77.1	412	2	Q7TN69	Q7tn69 hystrix bra
27	37	77.1	1581	2	Q6C494	Q6c494 yarrowia li
28	36	75.0	153	2	Q9FXM2	Q9fxm2 arabidopsis
29	36	75.0	202	2	Q8L8T0	Q8l8t0 arabidopsis
30	36	75.0	202	2	Q9LVA4	Q9lva4 arabidopsis
31	36	75.0	217	2	Q98E60	Q98e60 rhizobium l
32	36	75.0	430	2	Q7BUD9	Q7bud9 amycolatops
33	36	75.0	430	2	AAS07756	Aas07756 amycolato
34	36	75.0	432	2	O52565	O52565 amycolatops
35	35	72.9	288	2	Q8FD65	Q8fd65 escherichia
36	35	72.9	290	2	Q06258	Q06258 streptomyce
37	35	72.9	291	2	Q8VWD2	Q8vwd2 streptomyce
38	35	72.9	309	1	YHCC_ECOLI	P45476 escherichia
39	35	72.9	309	2	Q7UBF4	Q7ubf4 shigella fl
40	35	72.9	320	2	Q83JF2	Q83jf2 shigella fl
41	35	72.9	467	2	Q88C72	Q88c72 pseudomonas
42	35	72.9	562	2	Q88KJ7	Q88kj7 pseudomonas
43	35	72.9	705	1	DAS_CANBO	O93884 candida boi
44	35	72.9	914	2	Q6NAC6	Q6nac6 rhodopseudo
45	35	72.9	914	2	CAE26702	Cae26702 rhodopseu

ALIGNMENTS

RESULT 1
Q7YQR5
ID Q7YQR5 PRELIMINARY; PRT; 414 AA.
AC Q7YQR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein B 100 (Fragment).
GN Name=apoB-100;
OS Aotus vociferans (Spix's owl monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=57176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22761261; PubMed=12878460;
RA Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT "A new phylogenetic marker, apolipoprotein B, provides compelling
RT evidence for eutherian relationships.";
RL Mol. Phylogenet. Evol. 28:225-240(2003).
DR EMBL; AF548396; AAP97352.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45955 MW; EEFA8492157E1BDE CRC64;

Query Match 91.7%; Score 44; DB 2; Length 414;
Best Local Similarity 90.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db        258 TRLTRKRGLK 267
```

RESULT 2
Q28473
ID Q28473 PRELIMINARY; PRT; 596 AA.
AC Q28473;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Apolipoprotein B (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92075708; PubMed=1742325;
RA Pape M.E., Castle C.K., Murray R.W., Funk G.M., Hunt C.E.,
RA Marotti K.R., Melchior G.W.;

```
RT   "Apo B metabolism in the cynomolgus monkey: evidence for post-
RT   transcriptional regulation.";
RL   Biochim. Biophys. Acta 1086:326-334(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Murray R.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X15737; CAA33755.1; -.
DR   PIR; S32802; S32802.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     596    596
SQ   SEQUENCE   596 AA;  66757 MW;  B13BBA74E25C3120 CRC64;

  Query Match             91.7%;  Score 44;  DB 2;  Length 596;
  Best Local Similarity   90.0%;  Pred. No. 2.3;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db        226 TRLTRKRGLK 235


RESULT 3
Q13788
ID   Q13788      PRELIMINARY;      PRT;  3262 AA.
AC   Q13788;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   APOB protein (Fragment).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191999; PubMed=2883086;
RA   Carlsson P., Darnfors C., Olofsson S.O., Bjursell G.;
RT   "Analysis of the human apolipoprotein B gene; complete structure of
RT   the B-74 region.";
RL   Gene 49:29-51(1986).
DR   EMBL; M15421; AAA51758.1; -.
DR   PIR; A27850; LPHUB.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005319; F:lipid transporter activity; NAS.
DR   GO; GO:0006869; P:lipid transport; NAS.
FT   NON_TER       1      1
SQ   SEQUENCE   3262 AA;  370140 MW;  56603BC0618DD40D CRC64;

  Query Match             91.7%;  Score 44;  DB 2;  Length 3262;
  Best Local Similarity   90.0%;  Pred. No. 15;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db       2084 TRLTRKRGLK 2093


RESULT 4
APB_HUMAN
ID   APB_HUMAN      STANDARD;      PRT;  4563 AA.
AC   P04114; O00502; Q13787;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein
DE   B-48 (Apo B-48)].
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87016385; PubMed=3763409;
RA   Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.,
RA   Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
RT   "Complete cDNA and derived protein sequence of human apolipoprotein B-
RT   100.";
RL   Nucleic Acids Res. 14:7501-7503(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=88003974; PubMed=3652907;
RA   Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
RA   Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
RT   "DNA sequence of the human apolipoprotein B gene.";
RL   DNA 6:363-372(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS ILE-98 AND GLU-4181.
RX   MEDLINE=87008488; PubMed=3759943;
RA   Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
RA   Gotto A.M. Jr., Chan L.;
RT   "The complete cDNA and amino acid sequence of human apolipoprotein B-
RT   100.";
RL   J. Biol. Chem. 261:12918-12921(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87041416; PubMed=3464946;
RA   Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
RA   Lee N., Brewer H.B. Jr.;
RT   "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
RT   derived amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87161758; PubMed=3030729;
RA   Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
RA   Zannis V.I.;
RT   "The complete sequence and structural analysis of human apolipoprotein
RT   B-100: relationship between apoB-100 and apoB-48 forms.";
RL   EMBO J. 5:3495-3507(1986).
RN   [6]
RP   SEQUENCE OF 709-906 FROM N.A.
RX   MEDLINE=85270450; PubMed=3860836;
RA   Deeb S.S., Motulsky A.G., Albers J.J.;
RT   "A partial cDNA clone for human apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
RN   [7]
RP   SEQUENCE OF 3056-3159 FROM N.A.
RX   MEDLINE=86041888; PubMed=3903660;
RA   Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
RA   Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
RT   "Human apolipoprotein B: identification of cDNA clones and
RT   characterization of mRNA.";
RL   Nucleic Acids Res. 13:6937-6953(1985).
RN   [8]
RP   SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=86093680; PubMed=3841204;
RA   Carlsson P., Olofsson S.O., Bondjers G., Darnfors C., Wiklund O.,
RA   Bjursell G.;
RT   "Molecular cloning of human apolipoprotein B cDNA.";
RL   Nucleic Acids Res. 13:8813-8826(1985).
RN   [9]
RP   SEQUENCE OF 3109-4563 FROM N.A.
RX   MEDLINE=85300528; PubMed=2994225;
RA   Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F., Urdea M.S.,
RA   Levy-Wilson B., Powell L.M., Pease R.J., Eddy R., Nakai H., Byers M.,
RA   Priestley L.M., Robertson E., Rall L.B., Betsholtz C., Shows T.B.,
RA   Mahley R.W., Scott J.;
RT   "Human apolipoprotein B: structure of carboxyl-terminal domains, sites
RT   of gene expression, and chromosomal localization.";
RL   Science 230:37-43(1985).
RN   [10]
```

```
RP   SEQUENCE OF 1-291 FROM N.A.
RX   MEDLINE=86149325; PubMed=3513177;
RA   Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
RA   Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
RT   "Isolation of a cDNA clone encoding the amino-terminal region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
RN   [11]
RP   SEQUENCE OF 1-1670 FROM N.A., AND VARIANT ILE-98.
RX   MEDLINE=86287319; PubMed=3461454;
RA   Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W., Yamanaka M.,
RA   Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
RT   "Analysis of cDNA clones encoding the entire B-26 region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
RN   [12]
RP   PARTIAL SEQUENCE, AND IDENTIFICATION OF APO-B48.
RX   MEDLINE=88018019; PubMed=3659919;
RA   Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
RA   Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
RA   Gotto A.M. Jr., Li W.-H., Chan L.;
RT   "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
RT   specific in-frame stop codon.";
RL   Science 238:363-366(1987).
RN   [13]
RP   DOMAINS.
RX   MEDLINE=87039351; PubMed=3773997;
RA   Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
RA   Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
RA   Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
RA   Levy-Wilson B., Scott J.;
RT   "Complete protein sequence and identification of structural domains of
RT   human apolipoprotein B.";
RL   Nature 323:734-738(1986).
RN   [14]
RP   DOMAINS.
RA   Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
RA   Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
RA   Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
RT   "Sequence, structure, receptor-binding domains and internal repeats of
RT   human apolipoprotein B-100.";
RL   Nature 323:738-742(1986).
RN   [15]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=86242245; PubMed=3087360;
RA   Dashti N., Lee D.M., Mok T.;
RT   "Apolipoprotein B is a calcium binding protein.";
RL   Biochem. Biophys. Res. Commun. 137:493-499(1986).
RN   [16]
RP   PALMITOYLATION OF CYS-1112.
RX   MEDLINE=20143590; PubMed=10679026;
RA   Zhao Y., McCabe J.B., Vance J., Berthiaume L.G.;
RT   "Palmitoylation of apolipoprotein B is required for proper
RT   intracellular sorting and transport of cholesteroyl esters and
RT   triglycerides.";
RL   Mol. Biol. Cell 11:721-734(2000).
RN   [17]
RP   VARIANT SER-4338.
RX   MEDLINE=91071750; PubMed=1979313;
RA   Navajas M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
RA   Cunny G., Cambien F., Roizes G.;
RT   "Detection by denaturing gradient gel electrophoresis of a new
RT   polymorphism in the apolipoprotein B gene.";
RL   Hum. Genet. 86:91-93(1990).
RN   [18]
RP   VARIANT FDB GLN-3527.
RX   MEDLINE=89098975; PubMed=2563166;
RA   Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
RA   McCarthy B.J.;
RT   "Association between a specific apolipoprotein B mutation and familial
RT   defective apolipoprotein B-100.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
RN   [19]
RP   VARIANT LEU-2739.
RX   MEDLINE=91016974; PubMed=2216805;
RA   Huang L.-S., Gavish D., Breslow J.L.;
RT   "Sequence polymorphism in the human apoB gene at position 8344.";
RL   Nucleic Acids Res. 18:5922-5922(1990).
RN   [20]
RP   VARIANT FDB CYS-3558.
RX   MEDLINE=95190020; PubMed=7883971;
RA   Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
RA   Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
RT   "Familial ligand-defective apolipoprotein B. Identification of a new
RT   mutation that decreases LDL receptor binding affinity.";
RL   J. Clin. Invest. 95:1225-1234(1995).
RN   [21]
RP   VARIANTS LEU-1437; SER-1914; LYS-2566; THR-3121; ALA-3945; MET-4128
RP   AND THR-4481.
RX   MEDLINE=97044521; PubMed=8889592;
RA   Poirier O., Ricard S., Behague I., Souriau C., Evans A.E.,
RA   Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
RT   "Detection of new variants in the apolipoprotein B (Apo B) gene by
RT   PCR-SSCP.";
RL   Hum. Mutat. 8:282-285(1996).
RN   [22]
RP   VARIANTS FDB GLN-3527 AND CYS-3558.
RX   MEDLINE=97403938; PubMed=9259199;
RA   Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
RA   Krempf M., Giraudet P., Junien C., Boileau C.;
RT   "Familial ligand-defective apolipoprotein B-100: simultaneous
RT   detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French
RT   population.";
RL   Hum. Mutat. 10:160-163(1997).
RN   [23]
RP   VARIANTS SER-1914; ARG-1923; LEU-2739; ASP-3319; THR-3427; GLN-3432
RP   AND ILE-3921.
RX   MEDLINE=98141125; PubMed=9490296;
RA   Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
RT   "Screening for mutations of the apolipoprotein B gene causing
RT   hypocholesterolemia.";
RL   Hum. Genet. 102:44-49(1998).
CC   -!- FUNCTION: Apolipoprotein B is a major protein constituent of
CC       chylomicrons (apo B-48), LDL (apo B-100) and VLDL (apo B-100). Apo
CC       B-100 functions as a recognition signal for the cellular binding
CC       and internalization of LDL particles by the apoB/E receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.

  Query Match             91.7%;  Score 44;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 21;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 5
Q7Z600
ID   Q7Z600      PRELIMINARY;      PRT;  4563 AA.
AC   Q7Z600;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Including Ag(X) antigen).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY324608; AAP72970.1; -.
DR   GO; GO:0005319; F:lipid transporter activity; IEA.
DR   GO; GO:0006869; P:lipid transport; IEA.
DR   InterPro; IPR009454; DUF1081.
DR   InterPro; IPR001747; Lipid_transprt_N.
DR   Pfam; PF06448; DUF1081; 1.
DR   Pfam; PF01347; Vitellogenin_N; 1.
DR   SMART; SM00638; LPD_N; 1.
KW   Lipoprotein.
SQ   SEQUENCE   4563 AA;  515553 MW;  030B34167CEDC63C CRC64;

  Query Match             91.7%;  Score 44;  DB 2;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 21;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              ||||| ||||
Db       3385 TRLTRKRGLK 3394


RESULT 6
Q7TN68
ID   Q7TN68      PRELIMINARY;      PRT;   421 AA.
AC   Q7TN68;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Glaucomys volans (Southern flying squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Petauristinae;
OC   Glaucomys.
OX   NCBI_TaxID=64683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243379; AAP50767.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     421    421
SQ   SEQUENCE   421 AA;  46747 MW;  D47B77BD4F864FD1 CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 421;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        264 SRLTRKRGLK 273


RESULT 7
Q7YR10
ID   Q7YR10      PRELIMINARY;      PRT;   432 AA.
AC   Q7YR10;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243375; AAP50763.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     432    432
SQ   SEQUENCE   432 AA;  48171 MW;  F27B7AB39604732C CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 432;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        275 SRLTRKRGLK 284


RESULT 8
Q7YQM8
ID   Q7YQM8      PRELIMINARY;      PRT;   436 AA.
AC   Q7YQM8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Nyctimene albiventer (Common tube-nosed fruit bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Nyctimene.
OX   NCBI_TaxID=48988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548435; AAP97391.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     436    436
SQ   SEQUENCE   436 AA;  48717 MW;  1C4A7EAD72D2C629 CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 436;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        279 SRLTRKRGLK 288


RESULT 9
Q7YQM7
ID   Q7YQM7      PRELIMINARY;      PRT;   438 AA.
AC   Q7YQM7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Pteropus hypomelanus (Small flying fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Pteropus.
OX   NCBI_TaxID=9405;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
```

```
DR   EMBL; AF548436; AAP97392.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48734 MW;  2BD85BCBF4E2CC41 CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        281 SRLTRKRGLK 290


RESULT 10
Q7YR04
ID   Q7YR04      PRELIMINARY;      PRT;   438 AA.
AC   Q7YR04;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Rousettus amplexicaudatus (Common rousette).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Rousettus.
OX   NCBI_TaxID=58083;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243383; AAP50771.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48597 MW;  41C890DEAF95C872 CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        281 SRLTRKRGLK 290


RESULT 11
Q7YR08
ID   Q7YR08      PRELIMINARY;      PRT;   445 AA.
AC   Q7YR08;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Chaetophractus villosus (South American armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Chaetophractus.
OX   NCBI_TaxID=29080;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243378; AAP50766.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49564 MW;  2DA5DC3ED2F0FDD2 CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 12
Q7TN64
ID   Q7TN64      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN64;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Agouti paca (Paca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Agoutidae; Agouti.
OX   NCBI_TaxID=108852;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548417; AAP97373.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49721 MW;  34AF7ABE90F121EF CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 13
Q7TN71
ID   Q7TN71      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Hydrochoerus hydrochaeris (Capybara) (Carpincho).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Hydrochaeridae;
OC   Hydrochaeris.
OX   NCBI_TaxID=10149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243369; AAP50757.1; -.
DR   InterPro; IPR000871; Beta_lactamase_A.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49520 MW;  CBBA2DD53D7A18D2 CRC64;
```

```
  Query Match             83.3%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 14
Q7TN72
ID   Q7TN72      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN72;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Erethizon dorsatum (North American porcupine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Erethizontidae;
OC   Erethizon.
OX   NCBI_TaxID=34844;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243368; AAP50756.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49617 MW;  9572FE5F5E7625F2 CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        288 SRLTRKRGLK 297


RESULT 15
Q60536
ID   Q60536      PRELIMINARY;      PRT;   780 AA.
AC   Q60536;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hamster apolipoprotein (apoB) (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236327; PubMed=2332175;
RA   Smith T.J., Hautamaa D., Maeda N.;
RT   "Sequence of the putative low-density lipoprotein receptor-binding
RT   regions of apolipoprotein B in mouse and hamster.";
RL   Gene 87:309-310(1990).
DR   EMBL; M35187; AAA37059.1; -.
DR   PIR; C60950; C60950.
DR   PIR; JH0102; JH0102.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     780    780
SQ   SEQUENCE   780 AA;  86625 MW;  E371D1B2079D8F7E CRC64;

  Query Match             83.3%;  Score 40;  DB 2;  Length 780;
  Best Local Similarity   80.0%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRSRGLK 10
              :|||| ||||
Db        642 SRLTRKRGLK 651


Search completed: December 29, 2004, 12:37:32
Job time : 59.5202 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-6
Perfect score: 49
Sequence: 1 TRLTRQRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	49	100.0	10	2	AAY30687	Aay30687 Apo-B100
2	46	93.9	10	2	AAY30682	Aay30682 Apo-B100
3	45	91.8	11	2	AAW57205	Aaw57205 Apo B bin
4	45	91.8	13	2	AAW57207	Aaw57207 Apo B 100
5	45	91.8	15	2	AAW41261	Aaw41261 Apolipopr
6	45	91.8	15	2	AAW96892	Aaw96892 ApoB-100
7	45	91.8	20	6	ABJ37575	Abj37575 Heparin b
8	45	91.8	22	2	AAW57208	Aaw57208 Apo B 100
9	45	91.8	22	2	AAW57209	Aaw57209 Apo B 100
10	45	91.8	34	5	AAE14541	Aae14541 Human apo
11	45	91.8	36	2	AAW96876	Aaw96876 Nucleic a
12	45	91.8	37	2	AAW64587	Aaw64587 Human apo
13	45	91.8	51	2	AAW96845	Aaw96845 Nucleic a
14	45	91.8	343	4	ABB37687	Abb37687 Peptide #
15	45	91.8	343	4	ABG52504	Abg52504 Human liv
16	45	91.8	377	2	AAR72704	Aar72704 Human apo
17	45	91.8	377	2	AAR34031	Aar34031 Sequence
18	45	91.8	2463	8	ADJ57400	Adj57400 Human apo
19	45	91.8	3923	2	AAY31237	Aay31237 Human Apo
20	45	91.8	4536	2	AAW41262	Aaw41262 Apolipopr
21	45	91.8	4536	2	AAW96826	Aaw96826 Amino aci
22	45	91.8	4560	5	AAU98981	Aau98981 Human apo
23	45	91.8	4561	7	ADD48677	Add48677 Human Pro
24	45	91.8	4563	5	AAO15893	Aao15893 Human apo
25	45	91.8	4563	6	ABR40253	Abr40253 Human ali
26	45	91.8	4563	6	ABU79140	Abu79140 Apolipopr
27	45	91.8	4563	7	ADF43408	Adf43408 Apolipopr
28	45	91.8	4563	8	ADH18871	Adh18871 Human apo
29	45	91.8	4563	8	ADH18870	Adh18870 Human apo
30	45	91.8	4563	8	ADO33445	Ado33445 Human apo
31	45	91.8	4563	8	ADO33447	Ado33447 Human apo
32	45	91.8	4590	4	AAU33184	Aau33184 Novel hum
33	44	89.8	10	2	AAY30683	Aay30683 Apo-B100
34	44	89.8	10	2	AAY30686	Aay30686 Apo-B100
35	43	87.8	10	2	AAY30684	Aay30684 Apo-B100
36	43	87.8	10	2	AAY30685	Aay30685 Apo-B100
37	40	81.6	10	2	AAY30690	Aay30690 Apo-B100
38	40	81.6	10	2	AAY30692	Aay30692 Apo-B100
39	40	81.6	10	2	AAY30688	Aay30688 Apo-B100
40	40	81.6	11	2	AAW57206	Aaw57206 Apo B 100
41	40	81.6	11	2	AAW87717	Aaw87717 Analogue
42	40	81.6	11	5	AAE21732	Aae21732 BSMR effe
43	40	81.6	11	6	ABU07938	Abu07938 Apoprotei
44	40	81.6	11	7	ADF56451	Adf56451 Human apo
45	40	81.6	12	2	AAW41260	Aaw41260 Apolipopr

ALIGNMENTS

RESULT 1
AAY30687
ID AAY30687 standard; peptide; 10 AA.
XX
AC AAY30687;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

```
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 49;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.019;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              ||||||||||
Db          1 TRLTRQRGLK 10


RESULT 2
AAY30682
ID   AAY30682 standard; peptide; 10 AA.
XX
AC   AAY30682;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             93.9%;  Score 46;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.069;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          1 TRLTRERGLK 10


RESULT 3
AAW57205
ID   AAW57205 standard; peptide; 11 AA.
XX
AC   AAW57205;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B binding site peptide 2.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 12; Page 52; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B binding site
CC   peptide which can be used as a component of a non-naturally occurring,
CC   receptor-competent low density lipoprotein (LDL) particle of the present
CC   invention. The LDL particle comprises at least 1 peptide component that
CC   has at least 1 binding site for an apo B protein receptor and at least 1
CC   lipophilic substituent. Also described in the invention are peptides
CC   containing an apo B binding sequence with at least 70% identity with
CC   sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC   naturally occurring, receptor-competent LDL particles are useful as: (i)
CC   drug-targeting vectors for delivering anticancer drugs to cancer cells
CC   that express an apo B protein receptor, and (ii) additives for cell
CC   culture media especially as growth supplements. Non-naturally occurring,
CC   receptor-competent LDL particles do not require the complete apo B
CC   sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 11 AA;

  Query Match             91.8%;  Score 45;  DB 2;  Length 11;
  Best Local Similarity   90.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          2 TRLTRKRGLK 11
```

```
RESULT 4
AAW57207
ID   AAW57207 standard; peptide; 13 AA.
XX
AC   AAW57207;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide B.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 13 AA;

  Query Match             91.8%;  Score 45;  DB 2;  Length 13;
  Best Local Similarity   90.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          3 TRLTRKRGLK 12


RESULT 5
AAW41261
ID   AAW41261 standard; peptide; 15 AA.
XX
AC   AAW41261;
XX
DT   19-MAY-1998  (first entry)
XX
DE   Apolipoprotein B-100 fragment.
XX
KW   Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW   thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW   angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW   prothrombinase complex.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9743311-A1.
XX
PD   20-NOV-1997.
XX
PF   09-MAY-1997;   97WO-GB001255.
XX
PR   09-MAY-1996;   96GB-00009702.
XX
PA   (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI   Bruckdorfer KR,  Ettelaie C;
XX
DR   WPI; 1998-008798/01.
XX
PT   Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT   used for treating or preventing coagulation, inhibiting angiogenesis,
PT   cell differentiation and apoptosis.
XX
PS   Disclosure; Page 22; 60pp; English.
XX
CC   This sequence is an example of the peptide of the invention. It has the
CC   formula (I), or their variants with one or more internal deletions,
CC   insertions or substitutions, while retaining anti-coagulant properties of
CC   apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC   Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC   (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC   aa. Compositions containing the peptide are used for simultaneous,
CC   separate or sequential treatment of cancer, particularly to prevent
CC   metastatic spread. They are also used to inhibit thromboplastin-mediated
CC   processes, specifically to prevent or reduce blood coagulation (e.g.
CC   during or after surgery or in cases of heart attack, stroke etc.) and to
CC   inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC   which is active as such or as part of a 98-aa peptide, inhibits
CC   activation of the prothrombinase complex; and prevents activation of
CC   factor VII on the surface of thromboplastin and of platelets by thrombin.
CC   It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC   smaller than apoB-100, they act more quickly
XX
SQ   Sequence 15 AA;

  Query Match             91.8%;  Score 45;  DB 2;  Length 15;
  Best Local Similarity   90.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          1 TRLTRKRGLK 10


RESULT 6
AAW96892
ID   AAW96892 standard; peptide; 15 AA.
XX
AC   AAW96892;
XX
DT   22-APR-1999  (first entry)
XX
DE   ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
```

KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from
CC human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC component of very-low density lipoproteins (VLDL), intermediate density
CC lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC present sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 15 AA;

Query Match 91.8%; Score 45; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.15;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
||||||:||||
Db 6 TRLTRKRGLK 15

RESULT 7
ABJ37575
ID ABJ37575 standard; peptide; 20 AA.
XX
AC ABJ37575;
XX
DT 10-MAY-2003 (first entry)
XX
DE Heparin binding peptide sequence #28.
XX
KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS Unidentified.
XX
PN WO2003007689-A2.
XX
PD 30-JAN-2003.
XX
PF 22-JUL-2002; 2002WO-US023419.
XX
PR 20-JUL-2001; 2001US-0306726P.
XX
PA (ETHZ-) ETH ZUERICH.
PA (UYZU-) UNIV ZURICH.
XX
PI Hubbell JA, Schoenmakers R, Maynard HD;
XX
DR WPI; 2003-300420/29.
XX
PT Use of a ligand comprising of at least one sulfated or sulfonated amino
PT acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT retinopathy and hypoxia.
XX
PS Disclosure; Fig 2; 79pp; English.
XX
CC The invention relates to a novel ligand for binding a target biomolecule,
CC which comprises a peptide having at least one sulphated or sulphonated
CC amino acid and at least one amino acid chosen from neutral and positively
CC charged amino acids. The novel ligands can be used for the treatment of
CC e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC This sequence represents a heparin binding peptide relating to the
CC invention
XX
SQ Sequence 20 AA;

Query Match 91.8%; Score 45; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.2;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
||||||:||||
Db 7 TRLTRKRGLK 16

RESULT 8
AAW57208
ID AAW57208 standard; peptide; 22 AA.
XX
AC AAW57208;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide C.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
FT Modified-site 22
FT /note= "attached to cholesterol"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.

XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 22 AA;

Query Match 91.8%; Score 45; DB 2; Length 22;
Best Local Similarity 90.0%; Pred. No. 0.22;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          7 TRLTRKRGLK 16
```

RESULT 9
AAW57209
ID AAW57209 standard; peptide; 22 AA.
XX
AC AAW57209;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide D.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 22 AA;

Query Match 91.8%; Score 45; DB 2; Length 22;
Best Local Similarity 90.0%; Pred. No. 0.22;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          7 TRLTRKRGLK 16
```

RESULT 10
AAE14541
ID AAE14541 standard; peptide; 34 AA.
XX
AC AAE14541;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human apoB-100 derived peptide p62.
XX
KW Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW cardiovascular disease; coronary heart disease; pre-eclampsia;
KW non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW peptide p62.
XX
OS Homo sapiens.
XX
PN WO200206314-A2.
XX
PD 24-JAN-2002.
XX
PF 18-JUL-2001; 2001WO-GB003212.
XX
PR 18-JUL-2000; 2000GB-00017641.
XX
PA (ARKT-) ARK THERAPEUTICS LTD.
XX
PI Narvanen O, Yla-Herttuala S;
XX
DR WPI; 2002-179777/23.
XX
PT New peptide useful in enzyme immunoassays for detecting oxidized low
PT density lipoprotein which is a marker of coronary heart disease and other
PT cardiovascular diseases, has affinity for oxidized low density
PT lipoprotein.
XX
PS Claim 6; Page 5; 21pp; English.
XX
CC The invention relates to peptides having affinity for oxidised low
CC density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC is useful in an immunoassay to determine the presence, and optionally,
CC the amount of antibodies in a sample, having affinity for oxLDL.
CC Preferably immobilised peptide is useful for measuring the amount of
CC autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC from a patient for evaluating the risk of coronary heart diseases, other
CC cardiovascular diseases, and several other disorders such as
CC periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC endothelial dysfunction. The peptide of the invention is stable, can be
CC synthesised easily without the need to isolate proteins from a patient's

```
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             91.8%;  Score 45;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.33;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db         25 TRLTRKRGLK 34


RESULT 11
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Fig 12C; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 36 AA;

  Query Match             91.8%;  Score 45;  DB 2;  Length 36;
  Best Local Similarity   90.0%;  Pred. No. 0.35;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db         11 TRLTRKRGLK 20


RESULT 12
AAW64587
ID   AAW64587 standard; peptide; 37 AA.
XX
AC   AAW64587;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Human apolipoprotein peptide fragment #1.
XX
KW   Factor V; human; detection; protein function; blood coagulation; apo;
KW   fat metabolism; Leyden mutation; deep vein thrombosis; apolipoprotein;
KW   Alzheimer's disease; 5,10-methylenetetrahydrofolate reductase; prion;
KW   hypercysteinaemia; factor VII; cardiovascular disease; pathogen; virus.
XX
OS   Homo sapiens.
XX
PN   EP857973-A2.
XX
PD   12-AUG-1998.
XX
PF   12-JAN-1998;   98EP-00890007.
XX
PR   13-JAN-1997;   97AT-00000044.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Moritz B,  Kiessig S,  Lang H,  Schenk V;
XX
DR   WPI; 1998-416142/36.
XX
PT   Detecting or quantifying mutant protein in presence of wild-type protein
PT   comprises reaction with ligand - used to detect mutant blood coagulation
PT   factors or apolipoproteins for diagnosing risk of thrombosis.
XX
PS   Example 2; Page 9; 18pp; German.
XX
CC   AAW64587 and AAW64588 are fragments of human apolipoprotein which are
CC   used with Factor V protein fragments in a novel method to detect the
CC   presence of a mutated protein in a sample that may also contain the
CC   corresponding wild-type protein. The method is used to detect mutations
CC   that alter protein functions (either point mutation or small insertions
CC   or deletions), particularly in proteins involved in blood coagulation or
CC   metabolism of fat. Protein functions which are specially detectable are
CC   the Leyden mutation in factor V (associated with increased risk of deep
CC   vein thrombosis), mutations in apolipoprotein (apo) genes (certain
CC   alleles of apoE indicates increased risk of developing Alzheimer's
CC   disease), thermostable 5,10-methylenetetrahydrofolate reductase
CC   (associated with hypercysteinaemia and venous thrombosis) and factor VII
CC   mutations (associated with increased risk of cardiovascular disease). The
CC   method can also be applied to proteins from pathogens, e.g. viruses or
CC   prions. The method does not require complex apparatus for polymerase
CC   chain reactions, it is simple, standardisable and reliable and is
CC   particularly suited to routine screening. It also allows mutant protein
CC   in a sample to be quantified
XX
SQ   Sequence 37 AA;

  Query Match             91.8%;  Score 45;  DB 2;  Length 37;
  Best Local Similarity   90.0%;  Pred. No. 0.35;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db         11 TRLTRKRGLK 20


RESULT 13
AAW96845
ID   AAW96845 standard; peptide; 51 AA.
```

```
XX
AC   AAW96845;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Page 151; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 51 AA;

  Query Match             91.8%;  Score 45;  DB 2;  Length 51;
  Best Local Similarity   90.0%;  Pred. No. 0.48;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db          6 TRLTRKRGLK 15


RESULT 14
ABB37687
ID   ABB37687 standard; peptide; 343 AA.
XX
AC   ABB37687;
XX
DT   04-FEB-2002  (first entry)
XX
DE   Peptide #5193 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 27; SEQ ID NO 30322; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a peptide encoded by a single exon nucleic acid probe
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 343 AA;

  Query Match             91.8%;  Score 45;  DB 4;  Length 343;
  Best Local Similarity   90.0%;  Pred. No. 2.7;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db        169 TRLTRKRGLK 178


RESULT 15
ABG52504
ID   ABG52504 standard; peptide; 343 AA.
XX
AC   ABG52504;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver peptide, SEQ ID No 31152.
XX
KW   Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW   hypercholesterolaemia; coronary heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
```

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 31152; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 343 AA;

Query Match 91.8%; Score 45; DB 4; Length 343;
Best Local Similarity 90.0%; Pred. No. 2.7;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTRQRGLK 10
              |||||:||||
Db        169 TRLTRKRGLK 178
```

Search completed: December 29, 2004, 12:28:49
Job time : 62.0227 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 9.65909 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-6
Perfect score: 49
Sequence: 1 TRLTRQRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	91.8	596	2	S32802	apolipoprotein B -
2	45	91.8	4563	1	LPHUB	apolipoprotein B-1
3	41	83.7	269	2	C60950	apolipoprotein B-1
4	41	83.7	779	2	JH0102	apolipoprotein B -
5	40	81.6	309	1	E65112	hypothetical 34.6
6	40	81.6	309	2	E85985	hypothetical prote
7	40	81.6	309	2	B91140	hypothetical prote
8	39	79.6	275	2	E60950	apolipoprotein B-1
9	39	79.6	309	2	AH0906	conserved hypothet
10	38	77.6	484	2	S40051	starch synthase (E
11	37	75.5	412	2	E83061	hypothetical prote
12	35	71.4	274	2	A60950	apolipoprotein B-1
13	35	71.4	393	2	S48288	probable phosphopr
14	35	71.4	784	2	JH0101	apolipoprotein B-1
15	34	69.4	430	2	AC2737	dihydroorotase [im
16	34	69.4	430	2	A97518	dihydroorotase (dh
17	34	69.4	493	2	AB2103	cobyric acid synth
18	34	69.4	1025	2	H86250	hypothetical prote
19	34	69.4	1230	2	T07663	soluble starch syn
20	33	67.3	310	2	T01266	starch synthase DU
21	33	67.3	330	1	F69534	pyruvate formate-l
22	33	67.3	332	2	B75286	hypothetical prote
23	33	67.3	392	2	T05350	adenylate transloc
24	33	67.3	476	2	C64119	starch synthase (E
25	33	67.3	477	2	B95130	glycogen synthase
26	33	67.3	477	2	H98000	starch synthase (E
27	33	67.3	498	2	D97492	replicative DNA he
28	33	67.3	498	2	AC2710	replicative DNA he
29	33	67.3	501	2	AG3437	replicative DNA he
30	33	67.3	680	2	AB1875	hypothetical prote
31	33	67.3	1674	2	T01265	starch synthase DU
32	33	67.3	2279	2	T42531	acetyl-CoA carboxy
33	33	67.3	2280	2	T38906	acetyl-CoA carboxy
34	32	65.3	83	2	S78289	ribosomal protein
35	32	65.3	107	2	S12607	salivary glue prot
36	32	65.3	112	2	S33822	salivary glue prot
37	32	65.3	173	2	G87383	acetyltransferase,
38	32	65.3	341	2	AE2652	exopolyphosphatase
39	32	65.3	394	2	T46858	molybdenum cofacto
40	32	65.3	427	2	B95936	probable glycosylt
41	32	65.3	477	2	AI0995	starch synthase (E
42	32	65.3	486	2	E86130	mannonate oxidored
43	32	65.3	486	2	B91289	D-mannonate oxidor
44	32	65.3	486	2	S56548	fructuronate reduc
45	32	65.3	552	1	E57987	cytochrome c-type

ALIGNMENTS

RESULT 1
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchi
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional r
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g9301
C;Superfamily: apolipoprotein B

Query Match 91.8%; Score 45; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 0.43;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTRQRGLK 10
              |||||:||||
Db        226 TRLTRKRGLK 235
```

RESULT 2
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A2
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Sc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNI
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: r
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986

A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732

R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.,
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than on
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.,
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place o
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

  Query Match 91.8%; Score 45; DB 1; Length 4563;
  Best Local Similarity 90.0%; Pred. No. 2.9;
  Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
|||||:||||
Db 3385 TRLTRKRGLK 3394

RESULT 3
C60950
apolipoprotein B-100 - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: C60950
A;Molecule type: DNA
A;Residues: 1-269 <LAW>
A;Cross-references: UNIPROT:Q60537; UNIPROT:Q60536
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match 83.7%; Score 41; DB 2; Length 269;
  Best Local Similarity 80.0%; Pred. No. 1.3;
  Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
:||||:||||
Db 216 SRLTRKRGLK 225

RESULT 4
JH0102
apolipoprotein B - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0102
R;Smith, T.J.
submitted to GenBank, June 1990
A;Reference number: A38864
A;Accession: JH0102
A;Molecule type: DNA
A;Residues: 1-779 <SMI>
A;Cross-references: UNIPROT:Q60536; GB:M35187
A;Note: this is a revision to the sequence from reference JH0101
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Contents: annotation
A;Note: this sequence has been revised in reference A38864
C;Genetics:
A;Gene: apoB
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

  Query Match 83.7%; Score 41; DB 2; Length 779;
  Best Local Similarity 80.0%; Pred. No. 3.5;
  Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
:||||:||||
Db 642 SRLTRKRGLK 651

RESULT 5
E65112
hypothetical 34.6 kD protein in arcB-gltB intergenic region - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E65112
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E65112
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-309 <BLAT>
A;Cross-references: UNIPROT:P45476; GB:AE000400; GB:U00096; NID:g2367203; PIDN:AAC76243
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

  Query Match 81.6%; Score 40; DB 1; Length 309;
  Best Local Similarity 80.0%; Pred. No. 2.3;
  Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
|:| ||||||
Db 170 TQLARQRGLK 179

RESULT 6
E85985
hypothetical protein yhcC [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85985
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca

Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85985
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <STO>
A;Cross-references: UNIPROT:P45476; GB:AE005174; NID:g12517832; PIDN:AAG58345.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yhcC
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 81.6%; Score 40; DB 2; Length 309;
Best Local Similarity 80.0%; Pred. No. 2.3;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
|:| ||||||
Db 170 TQLARQRGLK 179

RESULT 7
B91140
hypothetical protein ECs4090 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B91140
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91140
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <HAY>
A;Cross-references: UNIPROT:P45476; GB:BA000007; PIDN:BAB37513.1; PID:g13363563; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4090
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 81.6%; Score 40; DB 2; Length 309;
Best Local Similarity 80.0%; Pred. No. 2.3;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
|:| ||||||
Db 170 TQLARQRGLK 179

RESULT 8
E60950
apolipoprotein B-100 - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: E60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: E60950
A;Molecule type: mRNA
A;Residues: 1-275 <LAW>
A;Cross-references: UNIPROT:Q7LZ77
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

Query Match 79.6%; Score 39; DB 2; Length 275;
Best Local Similarity 80.0%; Pred. No. 3.3;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
| |||:||||
Db 221 TSLTRKRGLK 230

RESULT 9
AH0906
conserved hypothetical protein STY3508 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0906
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0906
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07846.1; PID:g16504394; GSPDB:GN00176
C;Genetics:
A;Gene: STY3508
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 79.6%; Score 39; DB 2; Length 309;
Best Local Similarity 70.0%; Pred. No. 3.7;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
||: |:||||
Db 170 TRIARERGLK 179

RESULT 10
S40051
starch synthase (EC 2.4.1.21) glgA - Bacillus subtilis
N;Alternate names: starch (bacterial glycogen) synthase glgA
C;Species: Bacillus subtilis
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S40051; E69632; S36627
R;Kiel, J.A.K.W.; Boels, J.M.; Beldman, G.; Venema, G.
Mol. Microbiol. 11, 203-218, 1994
A;Title: Glycogen in Bacillus subtilis: molecular characterization of an operon encoding
A;Reference number: S40048; MUID:94195107; PMID:8145641
A;Accession: S40051
A;Molecule type: DNA
A;Residues: 1-484 <KIE>
A;Cross-references: UNIPROT:P39125; EMBL:Z25795; NID:g397487; PIDN:CAA81043.1; PID:g5808
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69632
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-484 <KUN>
A;Cross-references: GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB15073.1; PID:g2635579

A;Experimental source: strain 168
C;Genetics:
A;Gene: glgA
A;Start codon: TTG
C;Superfamily: starch synthase
C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

  Query Match 77.6%; Score 38; DB 2; Length 484;
  Best Local Similarity 77.8%; Pred. No. 8.9;
  Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGL 9
 ||||:|:||
Db 300 TRLTKQKGL 308

RESULT 11
E83061
hypothetical protein PA4677 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83061
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83061
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-412 <STO>
A;Cross-references: UNIPROT:Q9HVB8; GB:AE004882; GB:AE004091; NID:g9950939; PIDN:AAG0806
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4677

  Query Match 75.5%; Score 37; DB 2; Length 412;
  Best Local Similarity 87.5%; Pred. No. 12;
  Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRG 8
 |:||||
Db 88 TKLTRQRG 95

RESULT 12
A60950
apolipoprotein B-100 - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: A60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: A60950
A;Molecule type: mRNA
A;Residues: 1-274 <LAW>
A;Cross-references: UNIPROT:Q7M2U9
A;Note: authors translated the codon GAT for residue 155 as His
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match 71.4%; Score 35; DB 2; Length 274;
  Best Local Similarity 87.5%; Pred. No. 21;
  Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LTRQRGLK 10
 |||:||||
Db 223 LTRKRGLK 230

RESULT 13
S48288
probable phosphoprotein phosphatase (EC 3.1.3.16) YBR0921 - yeast (Saccharomyces cerevi
N;Alternate names: protein YBR0921; protein YBR125c
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S48288; S45993; S44703
R;Mannhaupt, G.; Stucka, R.; Ehnle, S.; Vetter, I.; Feldmann, H.
Yeast 10, 1363-1381, 1994
A;Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
A;Reference number: S48255; MUID:95208357; PMID:7900426
A;Accession: S48288
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-393 <MAN>
A;Cross-references: UNIPROT:P38089; EMBL:X78993; NID:g476045; PIDN:CAA55626.1; PID:g476
R;Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45927
A;Accession: S45993
A;Molecule type: DNA
A;Residues: 1-393 <FE2>
A;Cross-references: EMBL:Z35994; NID:g536408; PID:g536409; MIPS:YBR125c
C;Genetics:
A;Gene: SGD:PTC4
A;Cross-references: SGD:S0000329
A;Map position: 2R
C;Superfamily: human phosphoprotein phosphatase 1A
C;Keywords: phosphoric monoester hydrolase

  Query Match 71.4%; Score 35; DB 2; Length 393;
  Best Local Similarity 77.8%; Pred. No. 29;
  Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRQRGL 9
 ||| |:|||
Db 384 TRLERERGL 392

RESULT 14
JH0101
apolipoprotein B-100 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0101; S33128; D60950
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Accession: JH0101
A;Molecule type: DNA
A;Residues: 1-784 <SMI>
A;Cross-references: UNIPROT:Q61314; GB:M35186
R;Smith, T.; Hautamaa, D.; Maeda, N.
submitted to the EMBL Data Library, May 1989
A;Reference number: S33128
A;Accession: S33128
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531,'S',533-784 <SM2>
A;Cross-references: EMBL:X15191
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDI
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: D60950
A;Molecule type: mRNA
A;Residues: 427-531,'S',533-700 <LAW>
C;Genetics:
A;Gene: MGI:Apob
A;Cross-references: MGI:88052
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein

F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

Query Match 71.4%; Score 35; DB 2; Length 784;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRQRGLK 10
              :|| |:||||
Db        647 SRLMRKRGLK 656
```

RESULT 15
AC2737
dihydroorotase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AC2737
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2737
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <KUR>
A;Cross-references: UNIPROT:Q8UFU0; GB:AE008688; PIDN:AAL42313.1; PID:g17739715; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: pyrC
A;Map position: circular chromosome
C;Superfamily: Bacillus dihydroorotase; Bacillus dihydroorotase homology

Query Match 69.4%; Score 34; DB 2; Length 430;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 RLTRQRGLK 10
              |: ||||:|
Db        249 RIARQRGVK 257
```

Search completed: December 29, 2004, 12:39:04
Job time : 10.6591 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 58.4091 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-6
Perfect score: 49
Sequence: 1 TRLTRQRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	91.8	414	2	Q7YQR5	Q7ygr5 aotus vocif
2	45	91.8	596	2	Q28473	macaca fasc
3	45	91.8	3262	2	Q13788	Q13788 homo sapien
4	45	91.8	4563	1	APB_HUMAN	P04114 homo sapien
5	45	91.8	4563	2	Q7Z600	Q7z600 homo sapien
6	41	83.7	421	2	Q7TN68	Q7tn68 glaucomys v
7	41	83.7	432	2	Q7YR10	Q7yr10 diceros bic
8	41	83.7	436	2	Q7YQM8	Q7yqm8 nyctimene a
9	41	83.7	438	2	Q7YQM7	Q7yqm7 pteropus hy
10	41	83.7	438	2	Q7YR04	Q7yr04 rousettus a
11	41	83.7	445	2	Q7YR08	Q7yr08 chaetophrac
12	41	83.7	445	2	Q7TN64	Q7tn64 agouti paca
13	41	83.7	445	2	Q7TN71	Q7tn71 hydrochoeru
14	41	83.7	445	2	Q7TN72	Q7tn72 erethizon d
15	41	83.7	780	2	Q60536	Q60536 mesocricetu
16	41	83.7	780	2	Q60537	Q60537 mesocricetu
17	40	81.6	309	1	YHCC_ECOLI	P45476 escherichia
18	40	81.6	309	2	Q7UBF4	Q7ubf4 shigella fl
19	40	81.6	320	2	Q83JF2	Q83jf2 shigella fl
20	39	79.6	275	2	Q7LZ77	Q7lz77 gallus gall
21	39	79.6	309	2	Q8XFV9	Q8xfv9 salmonella
22	39	79.6	309	2	Q7CPN5	Q7cpn5 salmonella
23	39	79.6	387	2	Q7YQN2	Q7yqn2 phalanger o
24	39	79.6	400	2	Q7YQM9	Q7yqm9 ornithorhyn
25	39	79.6	405	2	Q7YQN0	Q7yqn0 tachyglossu
26	39	79.6	445	2	Q7TN70	Q7tn70 dinomys bra
27	38	77.6	99	2	Q7UMF6	Q7umf6 rhodopirell
28	38	77.6	407	2	Q7TN65	Q7tn65 atherurus a
29	38	77.6	412	2	Q7TN69	Q7tn69 hystrix bra
30	38	77.6	476	1	GLGA_BACAN	Q81k85 bacillus an
31	38	77.6	476	1	GLGA_BACCR	Q816g8 bacillus ce
32	38	77.6	476	2	Q72YJ6	Q72yj6 bacillus ce
33	38	77.6	476	2	AAS43926	Aas43926 bacillus
34	38	77.6	476	2	AAT34248	Aat34248 bacillus
35	38	77.6	484	1	GLGA_BACSU	P39125 bacillus su
36	38	77.6	498	2	Q6HC18	Q6hc18 bacillus th
37	37	75.5	153	2	Q9FXM2	Q9fxm2 arabidopsis
38	37	75.5	202	2	Q8L8T0	Q8l8t0 arabidopsis
39	37	75.5	202	2	Q9LVA4	Q9lva4 arabidopsis
40	37	75.5	288	2	Q8FD65	Q8fd65 escherichia
41	37	75.5	412	2	Q9HVB8	Q9hvb8 pseudomonas
42	36	73.5	198	2	Q939I5	Q939i5 klebsiella
43	36	73.5	233	2	O06528	O06528 lactobacill
44	36	73.5	318	2	Q7YQP1	Q7yqp1 crocuta cro
45	36	73.5	581	2	Q8JL02	Q8jl02 virus phich

ALIGNMENTS

RESULT 1
Q7YQR5
ID Q7YQR5 PRELIMINARY; PRT; 414 AA.
AC Q7YQR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein B 100 (Fragment).
GN Name=apoB-100;
OS Aotus vociferans (Spix's owl monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=57176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22761261; PubMed=12878460;
RA Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT "A new phylogenetic marker, apolipoprotein B, provides compelling
RT evidence for eutherian relationships.";
RL Mol. Phylogenet. Evol. 28:225-240(2003).
DR EMBL; AF548396; AAP97352.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45955 MW; EEFA8492157E1BDE CRC64;

Query Match 91.8%; Score 45; DB 2; Length 414;
Best Local Similarity 90.0%; Pred. No. 0.81;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRQRGLK 10
|||||:||||
Db 258 TRLTRKRGLK 267

RESULT 2
Q28473
ID Q28473 PRELIMINARY; PRT; 596 AA.
AC Q28473;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Apolipoprotein B (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92075708; PubMed=1742325;
RA Pape M.E., Castle C.K., Murray R.W., Funk G.M., Hunt C.E.,
RA Marotti K.R., Melchior G.W.;

```
RT   "Apo B metabolism in the cynomolgus monkey: evidence for post-
RT   transcriptional regulation.";
RL   Biochim. Biophys. Acta 1086:326-334(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Murray R.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X15737; CAA33755.1; -.
DR   PIR; S32802; S32802.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     596    596
SQ   SEQUENCE   596 AA;  66757 MW;  B13BBA74E25C3120 CRC64;

  Query Match             91.8%;  Score 45;  DB 2;  Length 596;
  Best Local Similarity   90.0%;  Pred. No. 1.2;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db        226 TRLTRKRGLK 235


RESULT 3
Q13788
ID   Q13788      PRELIMINARY;      PRT;  3262 AA.
AC   Q13788;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   APOB protein (Fragment).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191999; PubMed=2883086;
RA   Carlsson P., Darnfors C., Olofsson S.O., Bjursell G.;
RT   "Analysis of the human apolipoprotein B gene; complete structure of
RT   the B-74 region.";
RL   Gene 49:29-51(1986).
DR   EMBL; M15421; AAA51758.1; -.
DR   PIR; A27850; LPHUB.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005319; F:lipid transporter activity; NAS.
DR   GO; GO:0006869; P:lipid transport; NAS.
FT   NON_TER       1      1
SQ   SEQUENCE   3262 AA;  370140 MW;  56603BC0618DD40D CRC64;

  Query Match             91.8%;  Score 45;  DB 2;  Length 3262;
  Best Local Similarity   90.0%;  Pred. No. 7.8;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db       2084 TRLTRKRGLK 2093


RESULT 4
APB_HUMAN
ID   APB_HUMAN      STANDARD;      PRT;  4563 AA.
AC   P04114; O00502; Q13787;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein
DE   B-48 (Apo B-48)].
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87016385; PubMed=3763409;
RA   Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.,
RA   Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
RT   "Complete cDNA and derived protein sequence of human apolipoprotein B-
RT   100.";
RL   Nucleic Acids Res. 14:7501-7503(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=88003974; PubMed=3652907;
RA   Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
RA   Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
RT   "DNA sequence of the human apolipoprotein B gene.";
RL   DNA 6:363-372(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS ILE-98 AND GLU-4181.
RX   MEDLINE=87008488; PubMed=3759943;
RA   Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
RA   Gotto A.M. Jr., Chan L.;
RT   "The complete cDNA and amino acid sequence of human apolipoprotein B-
RT   100.";
RL   J. Biol. Chem. 261:12918-12921(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87041416; PubMed=3464946;
RA   Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
RA   Lee N., Brewer H.B. Jr.;
RT   "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
RT   derived amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87161758; PubMed=3030729;
RA   Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
RA   Zannis V.I.;
RT   "The complete sequence and structural analysis of human apolipoprotein
RT   B-100: relationship between apoB-100 and apoB-48 forms.";
RL   EMBO J. 5:3495-3507(1986).
RN   [6]
RP   SEQUENCE OF 709-906 FROM N.A.
RX   MEDLINE=85270450; PubMed=3860836;
RA   Deeb S.S., Motulsky A.G., Albers J.J.;
RT   "A partial cDNA clone for human apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
RN   [7]
RP   SEQUENCE OF 3056-3159 FROM N.A.
RX   MEDLINE=86041888; PubMed=3903660;
RA   Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
RA   Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
RT   "Human apolipoprotein B: identification of cDNA clones and
RT   characterization of mRNA.";
RL   Nucleic Acids Res. 13:6937-6953(1985).
RN   [8]
RP   SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=86093680; PubMed=3841204;
RA   Carlsson P., Olofsson S.O., Bondjers G., Darnfors C., Wiklund O.,
RA   Bjursell G.;
RT   "Molecular cloning of human apolipoprotein B cDNA.";
RL   Nucleic Acids Res. 13:8813-8826(1985).
RN   [9]
RP   SEQUENCE OF 3109-4563 FROM N.A.
RX   MEDLINE=85300528; PubMed=2994225;
RA   Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F., Urdea M.S.,
RA   Levy-Wilson B., Powell L.M., Pease R.J., Eddy R., Nakai H., Byers M.,
RA   Priestley L.M., Robertson E., Rall L.B., Betsholtz C., Shows T.B.,
RA   Mahley R.W., Scott J.;
RT   "Human apolipoprotein B: structure of carboxyl-terminal domains, sites
RT   of gene expression, and chromosomal localization.";
RL   Science 230:37-43(1985).
RN   [10]
```

```
RP   SEQUENCE OF 1-291 FROM N.A.
RX   MEDLINE=86149325; PubMed=3513177;
RA   Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
RA   Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
RT   "Isolation of a cDNA clone encoding the amino-terminal region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
RN   [11]
RP   SEQUENCE OF 1-1670 FROM N.A., AND VARIANT ILE-98.
RX   MEDLINE=86287319; PubMed=3461454;
RA   Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W., Yamanaka M.,
RA   Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
RT   "Analysis of cDNA clones encoding the entire B-26 region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
RN   [12]
RP   PARTIAL SEQUENCE, AND IDENTIFICATION OF APO-B48.
RX   MEDLINE=88018019; PubMed=3659919;
RA   Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
RA   Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
RA   Gotto A.M. Jr., Li W.-H., Chan L.;
RT   "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
RT   specific in-frame stop codon.";
RL   Science 238:363-366(1987).
RN   [13]
RP   DOMAINS.
RX   MEDLINE=87039351; PubMed=3773997;
RA   Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
RA   Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
RA   Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
RA   Levy-Wilson B., Scott J.;
RT   "Complete protein sequence and identification of structural domains of
RT   human apolipoprotein B.";
RL   Nature 323:734-738(1986).
RN   [14]
RP   DOMAINS.
RA   Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
RA   Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
RA   Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
RT   "Sequence, structure, receptor-binding domains and internal repeats of
RT   human apolipoprotein B-100.";
RL   Nature 323:738-742(1986).
RN   [15]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=86242245; PubMed=3087360;
RA   Dashti N., Lee D.M., Mok T.;
RT   "Apolipoprotein B is a calcium binding protein.";
RL   Biochem. Biophys. Res. Commun. 137:493-499(1986).
RN   [16]
RP   PALMITOYLATION OF CYS-1112.
RX   MEDLINE=20143590; PubMed=10679026;
RA   Zhao Y., McCabe J.B., Vance J., Berthiaume L.G.;
RT   "Palmitoylation of apolipoprotein B is required for proper
RT   intracellular sorting and transport of cholesteroyl esters and
RT   triglycerides.";
RL   Mol. Biol. Cell 11:721-734(2000).
RN   [17]
RP   VARIANT SER-4338.
RX   MEDLINE=91071750; PubMed=1979313;
RA   Navajas M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
RA   Cunny G., Cambien F., Roizes G.;
RT   "Detection by denaturing gradient gel electrophoresis of a new
RT   polymorphism in the apolipoprotein B gene.";
RL   Hum. Genet. 86:91-93(1990).
RN   [18]
RP   VARIANT FDB GLN-3527.
RX   MEDLINE=89098975; PubMed=2563166;
RA   Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
RA   McCarthy B.J.;
RT   "Association between a specific apolipoprotein B mutation and familial
RT   defective apolipoprotein B-100.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
RN   [19]
RP   VARIANT LEU-2739.
RX   MEDLINE=91016974; PubMed=2216805;
RA   Huang L.-S., Gavish D., Breslow J.L.;
RT   "Sequence polymorphism in the human apoB gene at position 8344.";
RL   Nucleic Acids Res. 18:5922-5922(1990).
RN   [20]
RP   VARIANT FDB CYS-3558.
RX   MEDLINE=95190020; PubMed=7883971;
RA   Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
RA   Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
RT   "Familial ligand-defective apolipoprotein B. Identification of a new
RT   mutation that decreases LDL receptor binding affinity.";
RL   J. Clin. Invest. 95:1225-1234(1995).
RN   [21]
RP   VARIANTS LEU-1437; SER-1914; LYS-2566; THR-3121; ALA-3945; MET-4128
RP   AND THR-4481.
RX   MEDLINE=97044521; PubMed=8889592;
RA   Poirier O., Ricard S., Behague I., Souriau C., Evans A.E.,
RA   Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
RT   "Detection of new variants in the apolipoprotein B (Apo B) gene by
RT   PCR-SSCP.";
RL   Hum. Mutat. 8:282-285(1996).
RN   [22]
RP   VARIANTS FDB GLN-3527 AND CYS-3558.
RX   MEDLINE=97403938; PubMed=9259199;
RA   Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
RA   Krempf M., Giraudet P., Junien C., Boileau C.;
RT   "Familial ligand-defective apolipoprotein B-100: simultaneous
RT   detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French
RT   population.";
RL   Hum. Mutat. 10:160-163(1997).
RN   [23]
RP   VARIANTS SER-1914; ARG-1923; LEU-2739; ASP-3319; THR-3427; GLN-3432
RP   AND ILE-3921.
RX   MEDLINE=98141125; PubMed=9490296;
RA   Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
RT   "Screening for mutations of the apolipoprotein B gene causing
RT   hypocholesterolemia.";
RL   Hum. Genet. 102:44-49(1998).
CC   -!- FUNCTION: Apolipoprotein B is a major protein constituent of
CC       chylomicrons (apo B-48), LDL (apo B-100) and VLDL (apo B-100). Apo
CC       B-100 functions as a recognition signal for the cellular binding
CC       and internalization of LDL particles by the apoB/E receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.

  Query Match             91.8%;  Score 45;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 11;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db       3385 TRLTRKRGLK 3394


RESULT 5
Q7Z600
ID   Q7Z600      PRELIMINARY;      PRT;  4563 AA.
AC   Q7Z600;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Including Ag(X) antigen).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY324608; AAP72970.1; -.
DR   GO; GO:0005319; F:lipid transporter activity; IEA.
DR   GO; GO:0006869; P:lipid transport; IEA.
DR   InterPro; IPR009454; DUF1081.
DR   InterPro; IPR001747; Lipid_transprt_N.
DR   Pfam; PF06448; DUF1081; 1.
DR   Pfam; PF01347; Vitellogenin_N; 1.
DR   SMART; SM00638; LPD_N; 1.
KW   Lipoprotein.
SQ   SEQUENCE   4563 AA;  515553 MW;  030B34167CEDC63C CRC64;

  Query Match             91.8%;  Score 45;  DB 2;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 11;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              |||||:||||
Db       3385 TRLTRKRGLK 3394


RESULT 6
Q7TN68
ID   Q7TN68      PRELIMINARY;      PRT;   421 AA.
AC   Q7TN68;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Glaucomys volans (Southern flying squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Petauristinae;
OC   Glaucomys.
OX   NCBI_TaxID=64683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243379; AAP50767.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     421    421
SQ   SEQUENCE   421 AA;  46747 MW;  D47B77BD4F864FD1 CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 421;
  Best Local Similarity   80.0%;  Pred. No. 5.8;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        264 SRLTRKRGLK 273


RESULT 7
Q7YR10
ID   Q7YR10      PRELIMINARY;      PRT;   432 AA.
AC   Q7YR10;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243375; AAP50763.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     432    432
SQ   SEQUENCE   432 AA;  48171 MW;  F27B7AB39604732C CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 432;
  Best Local Similarity   80.0%;  Pred. No. 6;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        275 SRLTRKRGLK 284


RESULT 8
Q7YQM8
ID   Q7YQM8      PRELIMINARY;      PRT;   436 AA.
AC   Q7YQM8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Nyctimene albiventer (Common tube-nosed fruit bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Nyctimene.
OX   NCBI_TaxID=48988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548435; AAP97391.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     436    436
SQ   SEQUENCE   436 AA;  48717 MW;  1C4A7EAD72D2C629 CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 436;
  Best Local Similarity   80.0%;  Pred. No. 6.1;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        279 SRLTRKRGLK 288


RESULT 9
Q7YQM7
ID   Q7YQM7      PRELIMINARY;      PRT;   438 AA.
AC   Q7YQM7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Pteropus hypomelanus (Small flying fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Pteropus.
OX   NCBI_TaxID=9405;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
```

```
DR   EMBL; AF548436; AAP97392.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48734 MW;  2BD85BCBF4E2CC41 CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 6.1;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        281 SRLTRKRGLK 290


RESULT 10
Q7YR04
ID   Q7YR04      PRELIMINARY;      PRT;   438 AA.
AC   Q7YR04;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Rousettus amplexicaudatus (Common rousette).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Rousettus.
OX   NCBI_TaxID=58083;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243383; AAP50771.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48597 MW;  41C890DEAF95C872 CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 6.1;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        281 SRLTRKRGLK 290


RESULT 11
Q7YR08
ID   Q7YR08      PRELIMINARY;      PRT;   445 AA.
AC   Q7YR08;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Chaetophractus villosus (South American armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Chaetophractus.
OX   NCBI_TaxID=29080;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243378; AAP50766.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49564 MW;  2DA5DC3ED2F0FDD2 CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 6.2;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        288 SRLTRKRGLK 297


RESULT 12
Q7TN64
ID   Q7TN64      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN64;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Agouti paca (Paca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Agoutidae; Agouti.
OX   NCBI_TaxID=108852;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548417; AAP97373.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49721 MW;  34AF7ABE90F121EF CRC64;

  Query Match             83.7%;  Score 41;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 6.2;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        288 SRLTRKRGLK 297


RESULT 13
Q7TN71
ID   Q7TN71      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Hydrochoerus hydrochaeris (Capybara) (Carpincho).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Hydrochaeridae;
OC   Hydrochaeris.
OX   NCBI_TaxID=10149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243369; AAP50757.1; -.
DR   InterPro; IPR000871; Beta_lactamase_A.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49520 MW;  CBBA2DD53D7A18D2 CRC64;
```

```
  Query Match             83.7%;  Score 41;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 6.2;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        288 SRLTRKRGLK 297
```

```
RESULT 14
Q7TN72
ID   Q7TN72      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN72;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Erethizon dorsatum (North American porcupine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Erethizontidae;
OC   Erethizon.
OX   NCBI_TaxID=34844;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243368; AAP50756.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49617 MW;  9572FE5F5E7625F2 CRC64;
```

```
  Query Match             83.7%;  Score 41;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 6.2;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        288 SRLTRKRGLK 297
```

```
RESULT 15
Q60536
ID   Q60536      PRELIMINARY;      PRT;   780 AA.
AC   Q60536;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hamster apolipoprotein (apoB) (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236327; PubMed=2332175;
RA   Smith T.J., Hautamaa D., Maeda N.;
RT   "Sequence of the putative low-density lipoprotein receptor-binding
RT   regions of apolipoprotein B in mouse and hamster.";
RL   Gene 87:309-310(1990).
DR   EMBL; M35187; AAA37059.1; -.
DR   PIR; C60950; C60950.
DR   PIR; JH0102; JH0102.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     780    780
SQ   SEQUENCE   780 AA;  86625 MW;  E371D1B2079D8F7E CRC64;
```

```
  Query Match             83.7%;  Score 41;  DB 2;  Length 780;
  Best Local Similarity   80.0%;  Pred. No. 12;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRQRGLK 10
              :||||:||||
Db        642 SRLTRKRGLK 651
```

```
Search completed: December 29, 2004, 12:37:33
Job time : 59.5202 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-7
Perfect score: 49
Sequence: 1 TRLTEKRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	49	100.0	10	2	AAY30688	Aay30688 Apo-B100
2	46	93.9	10	2	AAY30689	Aay30689 Apo-B100
3	44	89.8	11	2	AAW57205	Aaw57205 Apo B bin
4	44	89.8	13	2	AAW57207	Aaw57207 Apo B 100
5	44	89.8	15	2	AAW41261	Aaw41261 Apolipopr
6	44	89.8	15	2	AAW96892	Aaw96892 ApoB-100
7	44	89.8	20	6	ABJ37575	Abj37575 Heparin b
8	44	89.8	22	2	AAW57208	Aaw57208 Apo B 100
9	44	89.8	22	2	AAW57209	Aaw57209 Apo B 100
10	44	89.8	34	5	AAE14541	Aae14541 Human apo
11	44	89.8	36	2	AAW96876	Aaw96876 Nucleic a
12	44	89.8	37	2	AAW64587	Aaw64587 Human apo
13	44	89.8	51	2	AAW96845	Aaw96845 Nucleic a
14	44	89.8	343	4	ABB37687	Abb37687 Peptide #
15	44	89.8	343	4	ABG52504	Abg52504 Human liv
16	44	89.8	377	2	AAR72704	Aar72704 Human apo
17	44	89.8	377	2	AAR34031	Aar34031 Sequence
18	44	89.8	2463	8	ADJ57400	Adj57400 Human apo
19	44	89.8	3923	2	AAY31237	Aay31237 Human Apo
20	44	89.8	4536	2	AAW41262	Aaw41262 Apolipopr
21	44	89.8	4536	2	AAW96826	Aaw96826 Amino aci
22	44	89.8	4560	5	AAU98981	Aau98981 Human apo
23	44	89.8	4561	7	ADD48677	Add48677 Human Pro
24	44	89.8	4563	5	AAO15893	Aao15893 Human apo
25	44	89.8	4563	6	ABR40253	Abr40253 Human ali
26	44	89.8	4563	6	ABU79140	Abu79140 Apolipopr
27	44	89.8	4563	7	ADF43408	Adf43408 Apolipopr
28	44	89.8	4563	8	ADH18871	Adh18871 Human apo
29	44	89.8	4563	8	ADH18870	Adh18870 Human apo
30	44	89.8	4563	8	ADO33445	Ado33445 Human apo
31	44	89.8	4563	8	ADO33447	Ado33447 Human apo
32	44	89.8	4590	4	AAU33184	Aau33184 Novel hum
33	40	81.6	10	2	AAY30682	Aay30682 Apo-B100
34	40	81.6	10	2	AAY30687	Aay30687 Apo-B100
35	39	79.6	10	2	AAY30690	Aay30690 Apo-B100
36	39	79.6	10	2	AAY30692	Aay30692 Apo-B100
37	39	79.6	10	2	AAY30686	Aay30686 Apo-B100
38	39	79.6	11	2	AAW57206	Aaw57206 Apo B 100
39	39	79.6	11	2	AAW87717	Aaw87717 Analogue
40	39	79.6	11	5	AAE21732	Aae21732 BSMR effe
41	39	79.6	11	6	ABU07938	Abu07938 Apoprotei
42	39	79.6	11	7	ADF56451	Adf56451 Human apo
43	39	79.6	12	2	AAW41260	Aaw41260 Apolipopr
44	39	79.6	15	2	AAW22911	Aaw22911 Low densi
45	39	79.6	16	1	AAP92302	Aap92302 Immunosup

ALIGNMENTS

RESULT 1
AAY30688
ID AAY30688 standard; peptide; 10 AA.
XX
AC AAY30688;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
||||||||||
Db 1 TRLTEKRGLK 10

RESULT 2
AAY30689
ID AAY30689 standard; peptide; 10 AA.
XX
AC AAY30689;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 93.9%; Score 46; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.027;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
||||:|||||
Db 1 TRLTDKRGLK 10

RESULT 3
AAW57205
ID AAW57205 standard; peptide; 11 AA.
XX
AC AAW57205;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B binding site peptide 2.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 12; Page 52; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B binding site
CC peptide which can be used as a component of a non-naturally occurring,
CC receptor-competent low density lipoprotein (LDL) particle of the present
CC invention. The LDL particle comprises at least 1 peptide component that
CC has at least 1 binding site for an apo B protein receptor and at least 1
CC lipophilic substituent. Also described in the invention are peptides
CC containing an apo B binding sequence with at least 70% identity with
CC sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC naturally occurring, receptor-competent LDL particles are useful as: (i)
CC drug-targeting vectors for delivering anticancer drugs to cancer cells
CC that express an apo B protein receptor, and (ii) additives for cell
CC culture media especially as growth supplements. Non-naturally occurring,
CC receptor-competent LDL particles do not require the complete apo B
CC sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 11 AA;

Query Match 89.8%; Score 44; DB 2; Length 11;
Best Local Similarity 90.0%; Pred. No. 0.074;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
|||| |||||
Db 2 TRLTRKRGLK 11

```
RESULT 4
AAW57207
ID   AAW57207 standard; peptide; 13 AA.
XX
AC   AAW57207;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide B.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 13 AA;

  Query Match             89.8%;  Score 44;  DB 2;  Length 13;
  Best Local Similarity   90.0%;  Pred. No. 0.088;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db          3 TRLTRKRGLK 12


RESULT 5
AAW41261
ID   AAW41261 standard; peptide; 15 AA.
XX
AC   AAW41261;
XX
DT   19-MAY-1998  (first entry)
XX
DE   Apolipoprotein B-100 fragment.
XX
KW   Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW   thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW   angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW   prothrombinase complex.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9743311-A1.
XX
PD   20-NOV-1997.
XX
PF   09-MAY-1997;   97WO-GB001255.
XX
PR   09-MAY-1996;   96GB-00009702.
XX
PA   (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI   Bruckdorfer KR,  Ettelaie C;
XX
DR   WPI; 1998-008798/01.
XX
PT   Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT   used for treating or preventing coagulation, inhibiting angiogenesis,
PT   cell differentiation and apoptosis.
XX
PS   Disclosure; Page 22; 60pp; English.
XX
CC   This sequence is an example of the peptide of the invention. It has the
CC   formula (I), or their variants with one or more internal deletions,
CC   insertions or substitutions, while retaining anti-coagulant properties of
CC   apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC   Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC   (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC   aa. Compositions containing the peptide are used for simultaneous,
CC   separate or sequential treatment of cancer, particularly to prevent
CC   metastatic spread. They are also used to inhibit thromboplastin-mediated
CC   processes, specifically to prevent or reduce blood coagulation (e.g.
CC   during or after surgery or in cases of heart attack, stroke etc.) and to
CC   inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC   which is active as such or as part of a 98-aa peptide, inhibits
CC   activation of the prothrombinase complex; and prevents activation of
CC   factor VII on the surface of thromboplastin and of platelets by thrombin.
CC   It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC   smaller than apoB-100, they act more quickly
XX
SQ   Sequence 15 AA;

  Query Match             89.8%;  Score 44;  DB 2;  Length 15;
  Best Local Similarity   90.0%;  Pred. No. 0.1;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db          1 TRLTRKRGLK 10


RESULT 6
AAW96892
ID   AAW96892 standard; peptide; 15 AA.
XX
AC   AAW96892;
XX
DT   22-APR-1999  (first entry)
XX
DE   ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
```

KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from
CC human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC component of very-low density lipoproteins (VLDL), intermediate density
CC lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC present sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 15 AA;

Query Match 89.8%; Score 44; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
|||| |||||
Db 6 TRLTRKRGLK 15

RESULT 7
ABJ37575
ID ABJ37575 standard; peptide; 20 AA.
XX
AC ABJ37575;
XX
DT 10-MAY-2003 (first entry)
XX
DE Heparin binding peptide sequence #28.
XX
KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS Unidentified.
XX
PN WO2003007689-A2.
XX
PD 30-JAN-2003.
XX
PF 22-JUL-2002; 2002WO-US023419.
XX
PR 20-JUL-2001; 2001US-0306726P.
XX
PA (ETHZ-) ETH ZUERICH.
PA (UYZU-) UNIV ZURICH.
XX
PI Hubbell JA, Schoenmakers R, Maynard HD;
XX
DR WPI; 2003-300420/29.
XX
PT Use of a ligand comprising of at least one sulfated or sulfonated amino
PT acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT retinopathy and hypoxia.
XX
PS Disclosure; Fig 2; 79pp; English.
XX
CC The invention relates to a novel ligand for binding a target biomolecule,
CC which comprises a peptide having at least one sulphated or sulphonated
CC amino acid and at least one amino acid chosen from neutral and positively
CC charged amino acids. The novel ligands can be used for the treatment of
CC e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC This sequence represents a heparin binding peptide relating to the
CC invention
XX
SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
|||| |||||
Db 7 TRLTRKRGLK 16

RESULT 8
AAW57208
ID AAW57208 standard; peptide; 22 AA.
XX
AC AAW57208;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide C.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
FT Modified-site 22
FT /note= "attached to cholesterol"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.

XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 22 AA;

Query Match 89.8%; Score 44; DB 2; Length 22;
Best Local Similarity 90.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
|||| |||||
Db 7 TRLTRKRGLK 16

RESULT 9
AAW57209
ID AAW57209 standard; peptide; 22 AA.
XX
AC AAW57209;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide D.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 22 AA;

Query Match 89.8%; Score 44; DB 2; Length 22;
Best Local Similarity 90.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
|||| |||||
Db 7 TRLTRKRGLK 16

RESULT 10
AAE14541
ID AAE14541 standard; peptide; 34 AA.
XX
AC AAE14541;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human apoB-100 derived peptide p62.
XX
KW Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW cardiovascular disease; coronary heart disease; pre-eclampsia;
KW non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW peptide p62.
XX
OS Homo sapiens.
XX
PN WO200206314-A2.
XX
PD 24-JAN-2002.
XX
PF 18-JUL-2001; 2001WO-GB003212.
XX
PR 18-JUL-2000; 2000GB-00017641.
XX
PA (ARKT-) ARK THERAPEUTICS LTD.
XX
PI Narvanen O, Yla-Herttuala S;
XX
DR WPI; 2002-179777/23.
XX
PT New peptide useful in enzyme immunoassays for detecting oxidized low
PT density lipoprotein which is a marker of coronary heart disease and other
PT cardiovascular diseases, has affinity for oxidized low density
PT lipoprotein.
XX
PS Claim 6; Page 5; 21pp; English.
XX
CC The invention relates to peptides having affinity for oxidised low
CC density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC is useful in an immunoassay to determine the presence, and optionally,
CC the amount of antibodies in a sample, having affinity for oxLDL.
CC Preferably immobilised peptide is useful for measuring the amount of
CC autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC from a patient for evaluating the risk of coronary heart diseases, other
CC cardiovascular diseases, and several other disorders such as
CC periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC endothelial dysfunction. The peptide of the invention is stable, can be
CC synthesised easily without the need to isolate proteins from a patient's

```
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             89.8%;  Score 44;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db         25 TRLTRKRGLK 34


RESULT 11
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Fig 12C; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 36 AA;

  Query Match             89.8%;  Score 44;  DB 2;  Length 36;
  Best Local Similarity   90.0%;  Pred. No. 0.25;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db         11 TRLTRKRGLK 20


RESULT 12
AAW64587
ID   AAW64587 standard; peptide; 37 AA.
XX
AC   AAW64587;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Human apolipoprotein peptide fragment #1.
XX
KW   Factor V; human; detection; protein function; blood coagulation; apo;
KW   fat metabolism; Leyden mutation; deep vein thrombosis; apolipoprotein;
KW   Alzheimer's disease; 5,10-methylenetetrahydrofolate reductase; prion;
KW   hypercysteinaemia; factor VII; cardiovascular disease; pathogen; virus.
XX
OS   Homo sapiens.
XX
PN   EP857973-A2.
XX
PD   12-AUG-1998.
XX
PF   12-JAN-1998;   98EP-00890007.
XX
PR   13-JAN-1997;   97AT-00000044.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Moritz B,  Kiessig S,  Lang H,  Schenk V;
XX
DR   WPI; 1998-416142/36.
XX
PT   Detecting or quantifying mutant protein in presence of wild-type protein
PT   comprises reaction with ligand - used to detect mutant blood coagulation
PT   factors or apolipoproteins for diagnosing risk of thrombosis.
XX
PS   Example 2; Page 9; 18pp; German.
XX
CC   AAW64587 and AAW64588 are fragments of human apolipoprotein which are
CC   used with Factor V protein fragments in a novel method to detect the
CC   presence of a mutated protein in a sample that may also contain the
CC   corresponding wild-type protein. The method is used to detect mutations
CC   that alter protein functions (either point mutation or small insertions
CC   or deletions), particularly in proteins involved in blood coagulation or
CC   metabolism of fat. Protein functions which are specially detectable are
CC   the Leyden mutation in factor V (associated with increased risk of deep
CC   vein thrombosis), mutations in apolipoprotein (apo) genes (certain
CC   alleles of apoE indicates increased risk of developing Alzheimer's
CC   disease), thermostable 5,10-methylenetetrahydrofolate reductase
CC   (associated with hypercysteinaemia and venous thrombosis) and factor VII
CC   mutations (associated with increased risk of cardiovascular disease). The
CC   method can also be applied to proteins from pathogens, e.g. viruses or
CC   prions. The method does not require complex apparatus for polymerase
CC   chain reactions, it is simple, standardisable and reliable and is
CC   particularly suited to routine screening. It also allows mutant protein
CC   in a sample to be quantified
XX
SQ   Sequence 37 AA;

  Query Match             89.8%;  Score 44;  DB 2;  Length 37;
  Best Local Similarity   90.0%;  Pred. No. 0.26;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db         11 TRLTRKRGLK 20


RESULT 13
AAW96845
ID   AAW96845 standard; peptide; 51 AA.
```

```
XX
AC   AAW96845;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Page 151; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 51 AA;

  Query Match             89.8%;  Score 44;  DB 2;  Length 51;
  Best Local Similarity   90.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db          6 TRLTRKRGLK 15


RESULT 14
ABB37687
ID   ABB37687 standard; peptide; 343 AA.
XX
AC   ABB37687;
XX
DT   04-FEB-2002  (first entry)
XX
DE   Peptide #5193 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 27; SEQ ID NO 30322; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a peptide encoded by a single exon nucleic acid probe
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 343 AA;

  Query Match             89.8%;  Score 44;  DB 4;  Length 343;
  Best Local Similarity   90.0%;  Pred. No. 2.6;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db        169 TRLTRKRGLK 178


RESULT 15
ABG52504
ID   ABG52504 standard; peptide; 343 AA.
XX
AC   ABG52504;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver peptide, SEQ ID No 31152.
XX
KW   Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW   hypercholesterolaemia; coronary heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
```

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 31152; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 343 AA;

Query Match 89.8%; Score 44; DB 4; Length 343;
Best Local Similarity 90.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTEKRGLK 10
              |||| |||||
Db        169 TRLTRKRGLK 178
```

Search completed: December 29, 2004, 12:28:49
Job time : 61.0227 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 9.65909 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-7
Perfect score: 49
Sequence: 1 TRLTEKRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	89.8	596	2	S32802	apolipoprotein B -
2	44	89.8	4563	1	LPHUB	apolipoprotein B-1
3	40	81.6	269	2	C60950	apolipoprotein B-1
4	40	81.6	779	2	JH0102	apolipoprotein B -
5	38	77.6	275	2	E60950	apolipoprotein B-1
6	36	73.5	406	2	AI0767	probable glycosylt
7	36	73.5	406	2	A90985	hypothetical prote
8	36	73.5	406	2	D85830	hypothetical prote
9	36	73.5	406	2	C64970	hypothetical prote
10	36	73.5	407	2	S52148	amsK protein - Erw
11	36	73.5	411	2	S15296	hypothetical prote
12	35	71.4	313	2	E69580	arabinan-endo 1,5-
13	35	71.4	461	2	S72953	probable GTP-bindi
14	35	71.4	462	2	F87080	probable GTP-bindi
15	34	69.4	198	2	T41529	hypothetical prote
16	34	69.4	232	2	E82104	hypothetical prote
17	34	69.4	258	2	T01873	hypothetical prote
18	34	69.4	274	2	A60950	apolipoprotein B-1
19	34	69.4	274	2	AG1306	D-alanyl-D-alanine
20	34	69.4	476	2	C64119	starch synthase (E
21	34	69.4	784	2	JH0101	apolipoprotein B-1
22	34	69.4	1073	2	T01955	hypothetical prote
23	34	69.4	1241	2	H84486	probable helicase
24	34	69.4	1265	2	F84517	probable helicase
25	34	69.4	1678	2	D86481	189.6K hypothetica
26	34	69.4	1752	2	T48965	hypothetical prote
27	33	67.3	210	2	I40540	vsrD protein - Pse
28	33	67.3	231	2	AF0336	aspartate racemase
29	33	67.3	309	2	AH0906	conserved hypothet
30	33	67.3	325	2	H69732	PBSX prophage ORF
31	33	67.3	436	2	F86486	protein F28J9.3 [i
32	33	67.3	484	2	S40051	starch synthase (E
33	33	67.3	490	2	C86486	protein F28J9.6 [i
34	33	67.3	493	2	E71008	hypothetical prote
35	33	67.3	559	2	T05129	hypothetical prote
36	33	67.3	676	2	A40363	DNA ligase (NAD) (
37	33	67.3	1058	2	S65460	apolipoprotein B -
38	32	65.3	125	1	ZTBPT9	gene 50 protein -
39	32	65.3	143	2	E97335	uncharacterized co
40	32	65.3	145	2	E90018	50S ribosomal prot
41	32	65.3	205	1	B64440	hypothetical prote
42	32	65.3	225	2	H70665	hypothetical prote
43	32	65.3	277	2	D82392	conserved hypothet
44	32	65.3	285	2	T27458	hypothetical prote
45	32	65.3	292	2	C83950	dipicolinate synth

ALIGNMENTS

RESULT 1
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchi
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional r
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g9301
C;Superfamily: apolipoprotein B

Query Match 89.8%; Score 44; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 0.86;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTEKRGLK 10
              |||| |||||
Db        226 TRLTRKRGLK 235
```

RESULT 2
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A2
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Sc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,'
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNI
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: r
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986

A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732

R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.;
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than on
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place o
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

```
A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

  Query Match             89.8%;  Score 44;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 5.7;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db       3385 TRLTRKRGLK 3394


RESULT 3
C60950
apolipoprotein B-100 - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: C60950
A;Molecule type: DNA
A;Residues: 1-269 <LAW>
A;Cross-references: UNIPROT:Q60537; UNIPROT:Q60536
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match             81.6%;  Score 40;  DB 2;  Length 269;
  Best Local Similarity   80.0%;  Pred. No. 2.6;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        216 SRLTRKRGLK 225


RESULT 4
JH0102
apolipoprotein B - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0102
R;Smith, T.J.
submitted to GenBank, June 1990
A;Reference number: A38864
A;Accession: JH0102
A;Molecule type: DNA
A;Residues: 1-779 <SMI>
A;Cross-references: UNIPROT:Q60536; GB:M35187
A;Note: this is a revision to the sequence from reference JH0101
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Contents: annotation
A;Note: this sequence has been revised in reference A38864
C;Genetics:
A;Gene: apoB
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

  Query Match             81.6%;  Score 40;  DB 2;  Length 779;
  Best Local Similarity   80.0%;  Pred. No. 7.1;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        642 SRLTRKRGLK 651


RESULT 5
E60950
apolipoprotein B-100 - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: E60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: E60950
A;Molecule type: mRNA
A;Residues: 1-275 <LAW>
A;Cross-references: UNIPROT:Q7LZ77
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein

  Query Match             77.6%;  Score 38;  DB 2;  Length 275;
  Best Local Similarity   80.0%;  Pred. No. 6.8;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              | || |||||
Db        221 TSLTRKRGLK 230


RESULT 6
AI0767
probable glycosyltransferase STY2310 [imported] - Salmonella enterica subsp. enterica s
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AI0767
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Parrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
```

```
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0767
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02463.1; PID:g16503330; GSPDB:GN00176
C;Genetics:
A;Gene: STY2310

  Query Match             73.5%;  Score 36;  DB 2;  Length 406;
  Best Local Similarity   87.5%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGL 9
              |||||:||
Db        230 RLTEKKGL 237


RESULT 7
A90985
hypothetical protein ECs2849 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A90985
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A90985
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <HAY>
A;Cross-references: UNIPROT:Q8X7P5; GB:BA000007; PIDN:BAB36272.1; PID:g13362317; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2849

  Query Match             73.5%;  Score 36;  DB 2;  Length 406;
  Best Local Similarity   87.5%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGL 9
              |||||:||
Db        230 RLTEKKGL 237


RESULT 8
D85830
hypothetical protein wcaL [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85830
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85830
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <STO>
A;Cross-references: UNIPROT:Q8X7P5; GB:AE005174; NID:g12516235; PIDN:AAG57104.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: wcaL

  Query Match             73.5%;  Score 36;  DB 2;  Length 406;
  Best Local Similarity   87.5%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGL 9
              |||||:||
Db        230 RLTEKKGL 237


RESULT 9
C64970
hypothetical protein b2044 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C64970
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C64970
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-406 <BLAT>
A;Cross-references: UNIPROT:P71243; GB:AE000295; GB:U00096; NID:g1788354; PIDN:AAC75105
A;Experimental source: strain K-12, substrain MG1655

  Query Match             73.5%;  Score 36;  DB 2;  Length 406;
  Best Local Similarity   87.5%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGL 9
              |||||:||
Db        230 RLTEKKGL 237


RESULT 10
S52148
amsK protein - Erwinia amylovora
C;Species: Erwinia amylovora
C;Date: 15-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S61901; S52148
R;Bugert, P.; Geider, K.
Mol. Microbiol. 15, 917-933, 1995
A;Title: Molecular analysis of the ams operon required for exopolysaccharide synthesis o
A;Reference number: S61891; MUID:95319333; PMID:7596293
A;Accession: S61901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <BU2>
A;Cross-references: UNIPROT:Q46638; EMBL:X77921; NID:g600426; PIDN:CAA54889.1; PID:g6004

  Query Match             73.5%;  Score 36;  DB 2;  Length 407;
  Best Local Similarity   87.5%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGL 9
              |||||:||
Db        231 RLTEKKGL 238


RESULT 11
S15296
hypothetical protein - Salmonella typhimurium
C;Species: Salmonella typhimurium
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S15296
R;Jiang, X.M.; Neal, B.; Santiago, F.; Lee, S.J.; Romana, L.K.; Reeves, P.R.
Mol. Microbiol. 5, 695-713, 1991
A;Title: Structure and sequence of the rfb (O antigen) gene cluster of Salmonella serova
A;Reference number: S15296; MUID:91260454; PMID:1710759
A;Accession: S15296
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <MOL>
A;Cross-references: UNIPROT:P26388
C;Keywords: transmembrane protein
```

```
  Query Match             73.5%;  Score 36;  DB 2;  Length 411;
  Best Local Similarity   87.5%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGL 9
              |||||:||
Db        235 RLTEKKGL 242


RESULT 12
E69580
arabinan-endo 1,5-alpha-L-arabinase abnA - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: E69580
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69580
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-313 <KUN>
A;Cross-references: UNIPROT:P94522; GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14841.
A;Experimental source: strain 168
C;Genetics:
A;Gene: abnA

  Query Match             71.4%;  Score 35;  DB 2;  Length 313;
  Best Local Similarity   70.0%;  Pred. No. 31;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              | |||:|||:
Db         59 TGLTEERGLR 68


RESULT 13
S72953
probable GTP-binding protein - Mycobacterium leprae
N;Alternate names: u0247e protein
C;Species: Mycobacterium leprae
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S72953
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A;Description: Mycobacterium leprae cosmid L247.
A;Reference number: S72589
A;Accession: S72953
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <SMI>
A;Cross-references: UNIPROT:Q49884; EMBL:U00021; NID:g467141; PIDN:AAA50911.1; PID:g4671
C;Genetics:
A;Start codon: GTG
C;Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
C;Keywords: duplication; GTP binding; nucleotide binding; P-loop
F;25-143/Domain: translation elongation factor Tu homology <ET1>
F;31-38/Region: nucleotide-binding motif A (P-loop)
F;140-143/Region: GTP-binding NKXD motif
F;152-154/Region: GTP-binding SAK/L motif
F;198-319/Domain: translation elongation factor Tu homology <ET2>
F;204-211/Region: nucleotide-binding motif A (P-loop)
F;316-319/Region: GTP-binding NKXD motif
F;349-351/Region: GTP-binding SAK/L motif

  Query Match             71.4%;  Score 35;  DB 2;  Length 461;
  Best Local Similarity   77.8%;  Pred. No. 44;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGLK 10
              |: ||||||
Db        451 RVREKRGLK 459


RESULT 14
F87080
probable GTP-binding protein [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: F87080
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: F87080
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <STO>
A;Cross-references: GB:AL450380; NID:g13093268; PIDN:CAC31753.1; GSPDB:GN00147
C;Genetics:
A;Gene: ML1372
C;Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongatio

  Query Match             71.4%;  Score 35;  DB 2;  Length 462;
  Best Local Similarity   77.8%;  Pred. No. 44;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGLK 10
              |: ||||||
Db        452 RVREKRGLK 460


RESULT 15
T41529
hypothetical protein SPCC645.12c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41529
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z22000
A;Accession: T41529
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-198 <WOO>
A;Cross-references: UNIPROT:Q9Y7V1; EMBL:AL049498; PIDN:CAB39908.1; GSPDB:GN00068; SPDB
A;Experimental source: strain 972h-; cosmid c645
C;Genetics:
A;Gene: SPDB:SPCC645.12c
A;Map position: 3

  Query Match             69.4%;  Score 34;  DB 2;  Length 198;
  Best Local Similarity   77.8%;  Pred. No. 32;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTEKRGLK 10
              || :|||||
Db         15 RLGQKRGLK 23
```

Search completed: December 29, 2004, 12:39:05
Job time : 10.6591 secs

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 58.4091 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-7
Perfect score: 49
Sequence: 1 TRLTEKRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	89.8	414	2	Q7YQR5	Q7yqr5 aotus vocif
2	44	89.8	596	2	Q28473	Q28473 macaca fasc
3	44	89.8	3262	2	Q13788	Q13788 homo sapien
4	44	89.8	4563	1	APB_HUMAN	P04114 homo sapien
5	44	89.8	4563	2	Q7Z600	Q7z600 homo sapien
6	40	81.6	421	2	Q7TN68	Q7tn68 glaucomys v
7	40	81.6	432	2	Q7YR10	Q7yr10 diceros bic
8	40	81.6	436	2	Q7YQM8	Q7yqm8 nyctimene a
9	40	81.6	438	2	Q7YQM7	Q7yqm7 pteropus hy
10	40	81.6	438	2	Q7YR04	Q7yr04 rousettus a
11	40	81.6	445	2	Q7YR08	Q7yr08 chaetophrac
12	40	81.6	445	2	Q7TN64	Q7tn64 agouti paca
13	40	81.6	445	2	Q7TN71	Q7tn71 hydrochoeru
14	40	81.6	445	2	Q7TN72	Q7tn72 erethizon d
15	40	81.6	780	2	Q60536	Q60536 mesocricetu
16	40	81.6	780	2	Q60537	Q60537 mesocricetu
17	38	77.6	275	2	Q7LZ77	Q7lz77 gallus gall
18	38	77.6	387	2	Q7YQN2	Q7yqn2 phalanger o
19	38	77.6	400	2	Q7YQM9	Q7yqm9 ornithorhyn
20	38	77.6	405	2	Q7YQN0	Q7yqn0 tachyglossu
21	38	77.6	445	2	Q7TN70	Q7tn70 dinomys bra
22	37	75.5	182	2	Q6L6Z2	Q616z2 thermoprote
23	37	75.5	182	2	BAD18908	Bad18908 thermopro
24	37	75.5	192	2	Q6L705	Q61705 thermoprote
25	37	75.5	192	2	BAD18895	Bad18895 thermopro
26	37	75.5	231	2	Q6ZG26	Q6zg26 oryza sativ
27	37	75.5	231	2	BAC98516	Bac98516 oryza sat
28	37	75.5	231	2	BAC98534	Bac98534 oryza sat
29	37	75.5	407	2	Q7TN65	Q7tn65 atherurus a
30	37	75.5	412	2	Q7TN69	Q7tn69 hystrix bra
31	37	75.5	845	2	Q6DCX0	Q6dcx0 xenopus lae
32	36	73.5	163	2	Q8S329	Q8s329 acetabulari
33	36	73.5	379	2	Q83KJ4	Q83kj4 shigella fl
34	36	73.5	404	2	Q7N1V8	Q7n1v8 photorhabdu
35	36	73.5	406	1	WCAL_ECOLI	P71243 escherichia
36	36	73.5	406	1	WCAL_SALTY	P26388 salmonella
37	36	73.5	406	2	Q7ACP5	Q7acp5 escherichia
38	36	73.5	406	2	Q7UCB6	Q7ucb6 shigella fl
39	36	73.5	406	2	Q8FG31	Q8fg31 escherichia
40	36	73.5	406	2	Q8Z5H7	Q8z5h7 salmonella
41	36	73.5	406	2	Q8X7P5	Q8x7p5 escherichia
42	36	73.5	407	1	AMSK_ERWAM	Q46638 erwinia amy
43	36	73.5	1108	2	Q9ARB8	Q9arb8 linum usita
44	35	71.4	112	1	PT17_STYPL	P28209 styela plic
45	35	71.4	132	2	Q8LCX0	Q8lcx0 arabidopsis

ALIGNMENTS

RESULT 1
Q7YQR5
ID Q7YQR5 PRELIMINARY; PRT; 414 AA.
AC Q7YQR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein B 100 (Fragment).
GN Name=apoB-100;
OS Aotus vociferans (Spix's owl monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=57176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22761261; PubMed=12878460;
RA Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT "A new phylogenetic marker, apolipoprotein B, provides compelling
RT evidence for eutherian relationships.";
RL Mol. Phylogenet. Evol. 28:225-240(2003).
DR EMBL; AF548396; AAP97352.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45955 MW; EEFA8492157E1BDE CRC64;

Query Match 89.8%; Score 44; DB 2; Length 414;
Best Local Similarity 90.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTEKRGLK 10
|||| |||||
Db 258 TRLTRKRGLK 267

RESULT 2
Q28473
ID Q28473 PRELIMINARY; PRT; 596 AA.
AC Q28473;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Apolipoprotein B (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92075708; PubMed=1742325;
RA Pape M.E., Castle C.K., Murray R.W., Funk G.M., Hunt C.E.,
RA Marotti K.R., Melchior G.W.;

```
RT   "Apo B metabolism in the cynomolgus monkey: evidence for post-
RT   transcriptional regulation.";
RL   Biochim. Biophys. Acta 1086:326-334(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Murray R.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X15737; CAA33755.1; -.
DR   PIR; S32802; S32802.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     596    596
SQ   SEQUENCE   596 AA;  66757 MW;  B13BBA74E25C3120 CRC64;

  Query Match             89.8%;  Score 44;  DB 2;  Length 596;
  Best Local Similarity   90.0%;  Pred. No. 3.7;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db        226 TRLTRKRGLK 235


RESULT 3
Q13788
ID   Q13788      PRELIMINARY;      PRT;  3262 AA.
AC   Q13788;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   APOB protein (Fragment).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191999; PubMed=2883086;
RA   Carlsson P., Darnfors C., Olofsson S.O., Bjursell G.;
RT   "Analysis of the human apolipoprotein B gene; complete structure of
RT   the B-74 region.";
RL   Gene 49:29-51(1986).
DR   EMBL; M15421; AAA51758.1; -.
DR   PIR; A27850; LPHUB.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005319; F:lipid transporter activity; NAS.
DR   GO; GO:0006869; P:lipid transport; NAS.
FT   NON_TER       1      1
SQ   SEQUENCE   3262 AA;  370140 MW;  56603BC0618DD40D CRC64;

  Query Match             89.8%;  Score 44;  DB 2;  Length 3262;
  Best Local Similarity   90.0%;  Pred. No. 20;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db       2084 TRLTRKRGLK 2093


RESULT 4
APB_HUMAN
ID   APB_HUMAN      STANDARD;      PRT;  4563 AA.
AC   P04114; O00502; Q13787;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein
DE   B-48 (Apo B-48)].
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87016385; PubMed=3763409;
RA   Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.,
RA   Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
RT   "Complete cDNA and derived protein sequence of human apolipoprotein B-
RT   100.";
RL   Nucleic Acids Res. 14:7501-7503(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=88003974; PubMed=3652907;
RA   Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
RA   Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
RT   "DNA sequence of the human apolipoprotein B gene.";
RL   DNA 6:363-372(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS ILE-98 AND GLU-4181.
RX   MEDLINE=87008488; PubMed=3759943;
RA   Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
RA   Gotto A.M. Jr., Chan L.;
RT   "The complete cDNA and amino acid sequence of human apolipoprotein B-
RT   100.";
RL   J. Biol. Chem. 261:12918-12921(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87041416; PubMed=3464946;
RA   Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
RA   Lee N., Brewer H.B. Jr.;
RT   "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
RT   derived amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87161758; PubMed=3030729;
RA   Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
RA   Zannis V.I.;
RT   "The complete sequence and structural analysis of human apolipoprotein
RT   B-100: relationship between apoB-100 and apoB-48 forms.";
RL   EMBO J. 5:3495-3507(1986).
RN   [6]
RP   SEQUENCE OF 709-906 FROM N.A.
RX   MEDLINE=85270450; PubMed=3860836;
RA   Deeb S.S., Motulsky A.G., Albers J.J.;
RT   "A partial cDNA clone for human apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
RN   [7]
RP   SEQUENCE OF 3056-3159 FROM N.A.
RX   MEDLINE=86041888; PubMed=3903660;
RA   Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
RA   Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
RT   "Human apolipoprotein B: identification of cDNA clones and
RT   characterization of mRNA.";
RL   Nucleic Acids Res. 13:6937-6953(1985).
RN   [8]
RP   SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=86093680; PubMed=3841204;
RA   Carlsson P., Olofsson S.O., Bondjers G., Darnfors C., Wiklund O.,
RA   Bjursell G.;
RT   "Molecular cloning of human apolipoprotein B cDNA.";
RL   Nucleic Acids Res. 13:8813-8826(1985).
RN   [9]
RP   SEQUENCE OF 3109-4563 FROM N.A.
RX   MEDLINE=85300528; PubMed=2994225;
RA   Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F., Urdea M.S.,
RA   Levy-Wilson B., Powell L.M., Pease R.J., Eddy R., Nakai H., Byers M.,
RA   Priestley L.M., Robertson E., Rall L.B., Betsholtz C., Shows T.B.,
RA   Mahley R.W., Scott J.;
RT   "Human apolipoprotein B: structure of carboxyl-terminal domains, sites
RT   of gene expression, and chromosomal localization.";
RL   Science 230:37-43(1985).
RN   [10]
```

```
RP   SEQUENCE OF 1-291 FROM N.A.
RX   MEDLINE=86149325; PubMed=3513177;
RA   Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
RA   Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
RT   "Isolation of a cDNA clone encoding the amino-terminal region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
RN   [11]
RP   SEQUENCE OF 1-1670 FROM N.A., AND VARIANT ILE-98.
RX   MEDLINE=86287319; PubMed=3461454;
RA   Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W., Yamanaka M.,
RA   Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
RT   "Analysis of cDNA clones encoding the entire B-26 region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
RN   [12]
RP   PARTIAL SEQUENCE, AND IDENTIFICATION OF APO-B48.
RX   MEDLINE=88018019; PubMed=3659919;
RA   Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
RA   Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
RA   Gotto A.M. Jr., Li W.-H., Chan L.;
RT   "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
RT   specific in-frame stop codon.";
RL   Science 238:363-366(1987).
RN   [13]
RP   DOMAINS.
RX   MEDLINE=87039351; PubMed=3773997;
RA   Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
RA   Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
RA   Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
RA   Levy-Wilson B., Scott J.;
RT   "Complete protein sequence and identification of structural domains of
RT   human apolipoprotein B.";
RL   Nature 323:734-738(1986).
RN   [14]
RP   DOMAINS.
RA   Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
RA   Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
RA   Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
RT   "Sequence, structure, receptor-binding domains and internal repeats of
RT   human apolipoprotein B-100.";
RL   Nature 323:738-742(1986).
RN   [15]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=86242245; PubMed=3087360;
RA   Dashti N., Lee D.M., Mok T.;
RT   "Apolipoprotein B is a calcium binding protein.";
RL   Biochem. Biophys. Res. Commun. 137:493-499(1986).
RN   [16]
RP   PALMITOYLATION OF CYS-1112.
RX   MEDLINE=20143590; PubMed=10679026;
RA   Zhao Y., McCabe J.B., Vance J., Berthiaume L.G.;
RT   "Palmitoylation of apolipoprotein B is required for proper
RT   intracellular sorting and transport of cholesteroyl esters and
RT   triglycerides.";
RL   Mol. Biol. Cell 11:721-734(2000).
RN   [17]
RP   VARIANT SER-4338.
RX   MEDLINE=91071750; PubMed=1979313;
RA   Navajas M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
RA   Cunny G., Cambien F., Roizes G.;
RT   "Detection by denaturing gradient gel electrophoresis of a new
RT   polymorphism in the apolipoprotein B gene.";
RL   Hum. Genet. 86:91-93(1990).
RN   [18]
RP   VARIANT FDB GLN-3527.
RX   MEDLINE=89098975; PubMed=2563166;
RA   Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
RA   McCarthy B.J.;
RT   "Association between a specific apolipoprotein B mutation and familial
RT   defective apolipoprotein B-100.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
RN   [19]
RP   VARIANT LEU-2739.
RX   MEDLINE=91016974; PubMed=2216805;
RA   Huang L.-S., Gavish D., Breslow J.L.;
RT   "Sequence polymorphism in the human apoB gene at position 8344.";
RL   Nucleic Acids Res. 18:5922-5922(1990).
RN   [20]
RP   VARIANT FDB CYS-3558.
RX   MEDLINE=95190020; PubMed=7883971;
RA   Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
RA   Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
RT   "Familial ligand-defective apolipoprotein B. Identification of a new
RT   mutation that decreases LDL receptor binding affinity.";
RL   J. Clin. Invest. 95:1225-1234(1995).
RN   [21]
RP   VARIANTS LEU-1437; SER-1914; LYS-2566; THR-3121; ALA-3945; MET-4128
RP   AND THR-4481.
RX   MEDLINE=97044521; PubMed=8889592;
RA   Poirier O., Ricard S., Behague I., Souriau C., Evans A.E.,
RA   Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
RT   "Detection of new variants in the apolipoprotein B (Apo B) gene by
RT   PCR-SSCP.";
RL   Hum. Mutat. 8:282-285(1996).
RN   [22]
RP   VARIANTS FDB GLN-3527 AND CYS-3558.
RX   MEDLINE=97403938; PubMed=9259199;
RA   Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
RA   Krempf M., Giraudet P., Junien C., Boileau C.;
RT   "Familial ligand-defective apolipoprotein B-100: simultaneous
RT   detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French
RT   population.";
RL   Hum. Mutat. 10:160-163(1997).
RN   [23]
RP   VARIANTS SER-1914; ARG-1923; LEU-2739; ASP-3319; THR-3427; GLN-3432
RP   AND ILE-3921.
RX   MEDLINE=98141125; PubMed=9490296;
RA   Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
RT   "Screening for mutations of the apolipoprotein B gene causing
RT   hypocholesterolemia.";
RL   Hum. Genet. 102:44-49(1998).
CC   -!- FUNCTION: Apolipoprotein B is a major protein constituent of
CC       chylomicrons (apo B-48), LDL (apo B-100) and VLDL (apo B-100). Apo
CC       B-100 functions as a recognition signal for the cellular binding
CC       and internalization of LDL particles by the apoB/E receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.

  Query Match             89.8%;  Score 44;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 28;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db       3385 TRLTRKRGLK 3394


RESULT 5
Q7Z600
ID   Q7Z600      PRELIMINARY;      PRT;  4563 AA.
AC   Q7Z600;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Including Ag(X) antigen).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY324608; AAP72970.1; -.
DR   GO; GO:0005319; F:lipid transporter activity; IEA.
DR   GO; GO:0006869; P:lipid transport; IEA.
DR   InterPro; IPR009454; DUF1081.
DR   InterPro; IPR001747; Lipid_transprt_N.
DR   Pfam; PF06448; DUF1081; 1.
DR   Pfam; PF01347; Vitellogenin_N; 1.
DR   SMART; SM00638; LPD_N; 1.
KW   Lipoprotein.
SQ   SEQUENCE   4563 AA;  515553 MW;  030B34167CEDC63C CRC64;

  Query Match             89.8%;  Score 44;  DB 2;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 28;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              |||| |||||
Db       3385 TRLTRKRGLK 3394


RESULT 6
Q7TN68
ID   Q7TN68      PRELIMINARY;      PRT;   421 AA.
AC   Q7TN68;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Glaucomys volans (Southern flying squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Petauristinae;
OC   Glaucomys.
OX   NCBI_TaxID=64683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243379; AAP50767.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     421    421
SQ   SEQUENCE   421 AA;  46747 MW;  D47B77BD4F864FD1 CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 421;
  Best Local Similarity   80.0%;  Pred. No. 17;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        264 SRLTRKRGLK 273


RESULT 7
Q7YR10
ID   Q7YR10      PRELIMINARY;      PRT;   432 AA.
AC   Q7YR10;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243375; AAP50763.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     432    432
SQ   SEQUENCE   432 AA;  48171 MW;  F27B7AB39604732C CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 432;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        275 SRLTRKRGLK 284


RESULT 8
Q7YQM8
ID   Q7YQM8      PRELIMINARY;      PRT;   436 AA.
AC   Q7YQM8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Nyctimene albiventer (Common tube-nosed fruit bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Nyctimene.
OX   NCBI_TaxID=48988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548435; AAP97391.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     436    436
SQ   SEQUENCE   436 AA;  48717 MW;  1C4A7EAD72D2C629 CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 436;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        279 SRLTRKRGLK 288


RESULT 9
Q7YQM7
ID   Q7YQM7      PRELIMINARY;      PRT;   438 AA.
AC   Q7YQM7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Pteropus hypomelanus (Small flying fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Pteropus.
OX   NCBI_TaxID=9405;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
```

```
DR   EMBL; AF548436; AAP97392.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48734 MW;  2BD85BCBF4E2CC41 CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        281 SRLTRKRGLK 290


RESULT 10
Q7YR04
ID   Q7YR04      PRELIMINARY;      PRT;   438 AA.
AC   Q7YR04;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Rousettus amplexicaudatus (Common rousette).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Rousettus.
OX   NCBI_TaxID=58083;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243383; AAP50771.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48597 MW;  41C890DEAF95C872 CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        281 SRLTRKRGLK 290


RESULT 11
Q7YR08
ID   Q7YR08      PRELIMINARY;      PRT;   445 AA.
AC   Q7YR08;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Chaetophractus villosus (South American armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Chaetophractus.
OX   NCBI_TaxID=29080;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243378; AAP50766.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49564 MW;  2DA5DC3ED2F0FDD2 CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


RESULT 12
Q7TN64
ID   Q7TN64      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN64;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Agouti paca (Paca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Agoutidae; Agouti.
OX   NCBI_TaxID=108852;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548417; AAP97373.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49721 MW;  34AF7ABE90F121EF CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


RESULT 13
Q7TN71
ID   Q7TN71      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Hydrochoerus hydrochaeris (Capybara) (Carpincho).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Hydrochaeridae;
OC   Hydrochaeris.
OX   NCBI_TaxID=10149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243369; AAP50757.1; -.
DR   InterPro; IPR000871; Beta_lactamase_A.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49520 MW;  CBBA2DD53D7A18D2 CRC64;
```

```
  Query Match             81.6%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


RESULT 14
Q7TN72
ID   Q7TN72      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN72;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Erethizon dorsatum (North American porcupine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Erethizontidae;
OC   Erethizon.
OX   NCBI_TaxID=34844;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243368; AAP50756.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49617 MW;  9572FE5F5E7625F2 CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 18;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


RESULT 15
Q60536
ID   Q60536      PRELIMINARY;      PRT;   780 AA.
AC   Q60536;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hamster apolipoprotein (apoB) (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236327; PubMed=2332175;
RA   Smith T.J., Hautamaa D., Maeda N.;
RT   "Sequence of the putative low-density lipoprotein receptor-binding
RT   regions of apolipoprotein B in mouse and hamster.";
RL   Gene 87:309-310(1990).
DR   EMBL; M35187; AAA37059.1; -.
DR   PIR; C60950; C60950.
DR   PIR; JH0102; JH0102.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     780    780
SQ   SEQUENCE   780 AA;  86625 MW;  E371D1B2079D8F7E CRC64;

  Query Match             81.6%;  Score 40;  DB 2;  Length 780;
  Best Local Similarity   80.0%;  Pred. No. 32;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTEKRGLK 10
              :||| |||||
Db        642 SRLTRKRGLK 651


Search completed: December 29, 2004, 12:37:34
Job time : 59.5202 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 61.0227 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-8
Perfect score: 50
Sequence: 1 TRLTDKRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	50	100.0	10	2	AAY30689	Aay30689 Apo-B100
2	46	92.0	10	2	AAY30688	Aay30688 Apo-B100
3	42	84.0	11	2	AAW57205	Aaw57205 Apo B bin
4	42	84.0	13	2	AAW57207	Aaw57207 Apo B 100
5	42	84.0	15	2	AAW41261	Aaw41261 Apolipopr
6	42	84.0	15	2	AAW96892	Aaw96892 ApoB-100
7	42	84.0	20	6	ABJ37575	Abj37575 Heparin b
8	42	84.0	22	2	AAW57208	Aaw57208 Apo B 100
9	42	84.0	22	2	AAW57209	Aaw57209 Apo B 100
10	42	84.0	34	5	AAE14541	Aae14541 Human apo
11	42	84.0	36	2	AAW96876	Aaw96876 Nucleic a
12	42	84.0	37	2	AAW64587	Aaw64587 Human apo
13	42	84.0	51	2	AAW96845	Aaw96845 Nucleic a
14	42	84.0	343	4	ABB37687	Abb37687 Peptide #
15	42	84.0	343	4	ABG52504	Abg52504 Human liv
16	42	84.0	377	2	AAR72704	Aar72704 Human apo
17	42	84.0	377	2	AAR34031	Aar34031 Sequence
18	42	84.0	2463	8	ADJ57400	Adj57400 Human apo
19	42	84.0	3923	2	AAY31237	Aay31237 Human Apo
20	42	84.0	4536	2	AAW41262	Aaw41262 Apolipopr
21	42	84.0	4536	2	AAW96826	Aaw96826 Amino aci
22	42	84.0	4560	5	AAU98981	Aau98981 Human apo
23	42	84.0	4561	7	ADD48677	Add48677 Human Pro
24	42	84.0	4563	5	AAO15893	Aao15893 Human apo
25	42	84.0	4563	6	ABR40253	Abr40253 Human ali
26	42	84.0	4563	6	ABU79140	Abu79140 Apolipopr
27	42	84.0	4563	7	ADF43408	Adf43408 Apolipopr
28	42	84.0	4563	8	ADH18871	Adh18871 Human apo
29	42	84.0	4563	8	ADH18870	Adh18870 Human apo
30	42	84.0	4563	8	ADO33445	Ado33445 Human apo
31	42	84.0	4563	8	ADO33447	Ado33447 Human apo
32	42	84.0	4590	4	AAU33184	Aau33184 Novel hum
33	39.5	79.0	11	2	AAY30699	Aay30699 Apo-B100
34	39	78.0	132	3	AAG57562	Aag57562 Arabidops
35	39	78.0	132	3	AAG59302	Aag59302 Arabidops
36	39	78.0	133	3	AAG39914	Aag39914 Arabidops
37	39	78.0	150	3	AAG57561	Aag57561 Arabidops
38	39	78.0	150	3	AAG59301	Aag59301 Arabidops
39	38	76.0	10	2	AAY30682	Aay30682 Apo-B100
40	38	76.0	10	2	AAY30687	Aay30687 Apo-B100
41	37	74.0	10	2	AAY30690	Aay30690 Apo-B100
42	37	74.0	10	2	AAY30692	Aay30692 Apo-B100
43	37	74.0	10	2	AAY30686	Aay30686 Apo-B100
44	37	74.0	11	2	AAW57206	Aaw57206 Apo B 100
45	37	74.0	11	2	AAW87717	Aaw87717 Analogue

ALIGNMENTS

RESULT 1
AAY30689
ID AAY30689 standard; peptide; 10 AA.
XX
AC AAY30689;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

```
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 50;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.0029;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              ||||||||||
Db          1 TRLTDKRGLK 10


RESULT 2
AAY30688
ID   AAY30688 standard; peptide; 10 AA.
XX
AC   AAY30688;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;

  Query Match             92.0%;  Score 46;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.019;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              ||||:|||||
Db          1 TRLTEKRGLK 10


RESULT 3
AAW57205
ID   AAW57205 standard; peptide; 11 AA.
XX
AC   AAW57205;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B binding site peptide 2.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 12; Page 52; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B binding site
CC   peptide which can be used as a component of a non-naturally occurring,
CC   receptor-competent low density lipoprotein (LDL) particle of the present
CC   invention. The LDL particle comprises at least 1 peptide component that
CC   has at least 1 binding site for an apo B protein receptor and at least 1
CC   lipophilic substituent. Also described in the invention are peptides
CC   containing an apo B binding sequence with at least 70% identity with
CC   sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC   naturally occurring, receptor-competent LDL particles are useful as: (i)
CC   drug-targeting vectors for delivering anticancer drugs to cancer cells
CC   that express an apo B protein receptor, and (ii) additives for cell
CC   culture media especially as growth supplements. Non-naturally occurring,
CC   receptor-competent LDL particles do not require the complete apo B
CC   sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 11 AA;

  Query Match             84.0%;  Score 42;  DB 2;  Length 11;
  Best Local Similarity   90.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db          2 TRLTRKRGLK 11
```

```
RESULT 4
AAW57207
ID   AAW57207 standard; peptide; 13 AA.
XX
AC   AAW57207;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Apo B 100 binding site peptide analogue peptide B.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 13 AA;

  Query Match             84.0%;  Score 42;  DB 2;  Length 13;
  Best Local Similarity   90.0%;  Pred. No. 0.16;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db          3 TRLTRKRGLK 12


RESULT 5
AAW41261
ID   AAW41261 standard; peptide; 15 AA.
XX
AC   AAW41261;
XX
DT   19-MAY-1998  (first entry)
XX
DE   Apolipoprotein B-100 fragment.
XX
KW   Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW   thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW   angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW   prothrombinase complex.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9743311-A1.
XX
PD   20-NOV-1997.
XX
PF   09-MAY-1997;   97WO-GB001255.
XX
PR   09-MAY-1996;   96GB-00009702.
XX
PA   (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI   Bruckdorfer KR,  Ettelaie C;
XX
DR   WPI; 1998-008798/01.
XX
PT   Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT   used for treating or preventing coagulation, inhibiting angiogenesis,
PT   cell differentiation and apoptosis.
XX
PS   Disclosure; Page 22; 60pp; English.
XX
CC   This sequence is an example of the peptide of the invention. It has the
CC   formula (I), or their variants with one or more internal deletions,
CC   insertions or substitutions, while retaining anti-coagulant properties of
CC   apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC   Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC   (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC   aa. Compositions containing the peptide are used for simultaneous,
CC   separate or sequential treatment of cancer, particularly to prevent
CC   metastatic spread. They are also used to inhibit thromboplastin-mediated
CC   processes, specifically to prevent or reduce blood coagulation (e.g.
CC   during or after surgery or in cases of heart attack, stroke etc.) and to
CC   inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC   which is active as such or as part of a 98-aa peptide, inhibits
CC   activation of the prothrombinase complex; and prevents activation of
CC   factor VII on the surface of thromboplastin and of platelets by thrombin.
CC   It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC   smaller than apoB-100, they act more quickly
XX
SQ   Sequence 15 AA;

  Query Match             84.0%;  Score 42;  DB 2;  Length 15;
  Best Local Similarity   90.0%;  Pred. No. 0.19;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db          1 TRLTRKRGLK 10


RESULT 6
AAW96892
ID   AAW96892 standard; peptide; 15 AA.
XX
AC   AAW96892;
XX
DT   22-APR-1999  (first entry)
XX
DE   ApoB-100 nuclear localisation signal sequence, residues 3353-3367.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
```

KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO9856938-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US011927.
XX
PR 13-JUN-1997; 97US-00874807.
PR 14-MAY-1998; 98US-00079030.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Guevara JG, Hoogeveen RC, Moore JP;
XX
DR WPI; 1999-070331/06.
XX
PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT used for delivering nucleic acid to cells for gene therapy and antisense
PT treatment.
XX
PS Claim 19; Fig 13D; 293pp; English.
XX
CC AAW96878-97 represent nuclear localisation signal sequence derived from
CC human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein
CC component of very-low density lipoproteins (VLDL), intermediate density
CC lipoprotein (IDL), low density lipoproteins ((LDL) and lipoprotein a. The
CC present sequence can be used in the composition of the invention. The
CC specification describes a composition that comprises LDL and
CC apolipoproteins for the binding and in vivo transport of nucleic acids.
CC The composition is used to deliver nucleic acids to eukaryotic cells, in
CC vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC molecule (or ribozyme). Specifically they are used for gene therapy of
CC cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC fibrosis and arteriosclerosis
XX
SQ Sequence 15 AA;

Query Match 84.0%; Score 42; DB 2; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTDKRGLK 10
|||| |||||
Db 6 TRLTRKRGLK 15

RESULT 7
ABJ37575
ID ABJ37575 standard; peptide; 20 AA.
XX
AC ABJ37575;
XX
DT 10-MAY-2003 (first entry)
XX
DE Heparin binding peptide sequence #28.
XX
KW Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW cardiovascular; circulatory; ligand; sulphated; sulphonated; tumour;
KW rheumatoid arthritis; hypoxia; diabetic retinopathy; heparin.
XX
OS Unidentified.
XX
PN WO2003007689-A2.
XX
PD 30-JAN-2003.
XX
PF 22-JUL-2002; 2002WO-US023419.
XX
PR 20-JUL-2001; 2001US-0306726P.
XX
PA (ETHZ-) ETH ZUERICH.
PA (UYZU-) UNIV ZURICH.
XX
PI Hubbell JA, Schoenmakers R, Maynard HD;
XX
DR WPI; 2003-300420/29.
XX
PT Use of a ligand comprising of at least one sulfated or sulfonated amino
PT acid for the treatment of e.g. tumors, rheumatoid arthritis, diabetic
PT retinopathy and hypoxia.
XX
PS Disclosure; Fig 2; 79pp; English.
XX
CC The invention relates to a novel ligand for binding a target biomolecule,
CC which comprises a peptide having at least one sulphated or sulphonated
CC amino acid and at least one amino acid chosen from neutral and positively
CC charged amino acids. The novel ligands can be used for the treatment of
CC e.g. tumours, rheumatoid arthritis, diabetic retinopathy and hypoxia.
CC This sequence represents a heparin binding peptide relating to the
CC invention
XX
SQ Sequence 20 AA;

Query Match 84.0%; Score 42; DB 6; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.26;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTDKRGLK 10
|||| |||||
Db 7 TRLTRKRGLK 16

RESULT 8
AAW57208
ID AAW57208 standard; peptide; 22 AA.
XX
AC AAW57208;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide C.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
FT Modified-site 22
FT /note= "attached to cholesterol"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.

XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 22 AA;

```
  Query Match             84.0%;  Score 42;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.28;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db          7 TRLTRKRGLK 16
```

RESULT 9
AAW57209
ID AAW57209 standard; peptide; 22 AA.
XX
AC AAW57209;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B 100 binding site peptide analogue peptide D.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "attached to retinoic acid"
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 13; Fig 7; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B 100 binding
CC site peptide analogue which can be used as a component of a non-
CC naturally occurring, receptor-competent low density lipoprotein (LDL)
CC particle of the present invention. The LDL particle comprises at least 1
CC peptide component that has at least 1 binding site for an apo B protein
CC receptor and at least 1 lipophilic substituent. Also described in the
CC invention are peptides containing an apo B binding sequence with at least
CC 70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC dimers. Non-naturally occurring, receptor-competent LDL particles are
CC useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC cancer cells that express an apo B protein receptor, and (ii) additives
CC for cell culture media especially as growth supplements. Non-naturally
CC occurring, receptor-competent LDL particles do not require the complete
CC apo B sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 22 AA;

```
  Query Match             84.0%;  Score 42;  DB 2;  Length 22;
  Best Local Similarity   90.0%;  Pred. No. 0.28;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db          7 TRLTRKRGLK 16
```

RESULT 10
AAE14541
ID AAE14541 standard; peptide; 34 AA.
XX
AC AAE14541;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human apoB-100 derived peptide p62.
XX
KW Oxidised low density lipoprotein; oxLDL; immunoassay; periaortitis;
KW cardiovascular disease; coronary heart disease; pre-eclampsia;
KW non-insulin-dependent diabetes; endothelial dysfunction; human; apoB-100;
KW peptide p62.
XX
OS Homo sapiens.
XX
PN WO200206314-A2.
XX
PD 24-JAN-2002.
XX
PF 18-JUL-2001; 2001WO-GB003212.
XX
PR 18-JUL-2000; 2000GB-00017641.
XX
PA (ARKT-) ARK THERAPEUTICS LTD.
XX
PI Narvanen O, Yla-Herttuala S;
XX
DR WPI; 2002-179777/23.
XX
PT New peptide useful in enzyme immunoassays for detecting oxidized low
PT density lipoprotein which is a marker of coronary heart disease and other
PT cardiovascular diseases, has affinity for oxidized low density
PT lipoprotein.
XX
PS Claim 6; Page 5; 21pp; English.
XX
CC The invention relates to peptides having affinity for oxidised low
CC density lipoprotein (oxLDL), in cyclised or multimeric form. The peptide
CC is useful in an immunoassay to determine the presence, and optionally,
CC the amount of antibodies in a sample, having affinity for oxLDL.
CC Preferably immobilised peptide is useful for measuring the amount of
CC autoantibodies for oxLDL in a sample, especially a serum or plasma sample
CC from a patient for evaluating the risk of coronary heart diseases, other
CC cardiovascular diseases, and several other disorders such as
CC periaortitis, pre-eclampsia, non-insulin-dependent diabetes and
CC endothelial dysfunction. The peptide of the invention is stable, can be
CC synthesised easily without the need to isolate proteins from a patient's

```
CC   blood, and has a long half-life. The present sequence is human apoB-100
CC   derived peptide p62 used in the invention
XX
SQ   Sequence 34 AA;

  Query Match             84.0%;  Score 42;  DB 5;  Length 34;
  Best Local Similarity   90.0%;  Pred. No. 0.45;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db         25 TRLTRKRGLK 34


RESULT 11
AAW96876
ID   AAW96876 standard; peptide; 36 AA.
XX
AC   AAW96876;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100, residues 3348-3390.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Fig 12C; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 36 AA;

  Query Match             84.0%;  Score 42;  DB 2;  Length 36;
  Best Local Similarity   90.0%;  Pred. No. 0.47;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db         11 TRLTRKRGLK 20


RESULT 12
AAW64587
ID   AAW64587 standard; peptide; 37 AA.
XX
AC   AAW64587;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Human apolipoprotein peptide fragment #1.
XX
KW   Factor V; human; detection; protein function; blood coagulation; apo;
KW   fat metabolism; Leyden mutation; deep vein thrombosis; apolipoprotein;
KW   Alzheimer's disease; 5,10-methylenetetrahydrofolate reductase; prion;
KW   hypercysteinaemia; factor VII; cardiovascular disease; pathogen; virus.
XX
OS   Homo sapiens.
XX
PN   EP857973-A2.
XX
PD   12-AUG-1998.
XX
PF   12-JAN-1998;   98EP-00890007.
XX
PR   13-JAN-1997;   97AT-00000044.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Moritz B,  Kiessig S,  Lang H,  Schenk V;
XX
DR   WPI; 1998-416142/36.
XX
PT   Detecting or quantifying mutant protein in presence of wild-type protein
PT   comprises reaction with ligand - used to detect mutant blood coagulation
PT   factors or apolipoproteins for diagnosing risk of thrombosis.
XX
PS   Example 2; Page 9; 18pp; German.
XX
CC   AAW64587 and AAW64588 are fragments of human apolipoprotein which are
CC   used with Factor V protein fragments in a novel method to detect the
CC   presence of a mutated protein in a sample that may also contain the
CC   corresponding wild-type protein. The method is used to detect mutations
CC   that alter protein functions (either point mutation or small insertions
CC   or deletions), particularly in proteins involved in blood coagulation or
CC   metabolism of fat. Protein functions which are specially detectable are
CC   the Leyden mutation in factor V (associated with increased risk of deep
CC   vein thrombosis), mutations in apolipoprotein (apo) genes (certain
CC   alleles of apoE indicates increased risk of developing Alzheimer's
CC   disease), thermostable 5,10-methylenetetrahydrofolate reductase
CC   (associated with hypercysteinaemia and venous thrombosis) and factor VII
CC   mutations (associated with increased risk of cardiovascular disease). The
CC   method can also be applied to proteins from pathogens, e.g. viruses or
CC   prions. The method does not require complex apparatus for polymerase
CC   chain reactions, it is simple, standardisable and reliable and is
CC   particularly suited to routine screening. It also allows mutant protein
CC   in a sample to be quantified
XX
SQ   Sequence 37 AA;

  Query Match             84.0%;  Score 42;  DB 2;  Length 37;
  Best Local Similarity   90.0%;  Pred. No. 0.49;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db         11 TRLTRKRGLK 20


RESULT 13
AAW96845
ID   AAW96845 standard; peptide; 51 AA.
```

```
XX
AC   AAW96845;
XX
DT   22-APR-1999  (first entry)
XX
DE   Nucleic acid binding domain from apoB-100.
XX
KW   Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;
KW   apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;
KW   nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;
KW   non-small cell lung carcinoma; diabetes; arteriosclerosis.
XX
OS   Homo sapiens.
XX
PN   WO9856938-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US011927.
XX
PR   13-JUN-1997;   97US-00874807.
PR   14-MAY-1998;   98US-00079030.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Guevara JG,  Hoogeveen RC,  Moore JP;
XX
DR   WPI; 1999-070331/06.
XX
PT   Composition comprising nucleic acid bound to LDL or VLDL lipoprotein -
PT   used for delivering nucleic acid to cells for gene therapy and antisense
PT   treatment.
XX
PS   Claim 16; Page 151; 293pp; English.
XX
CC   AAW96827-77 represent nucleic acid binding domains derived from human
CC   apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein component
CC   of very-low density lipoproteins (VLDL), intermediate density lipoprotein
CC   (IDL), low density lipoproteins ((LDL) and lipoprotein a. The present
CC   sequence can be used in the composition of the invention. The
CC   specification describes a composition that comprises LDL and
CC   apolipoproteins for the binding and in vivo transport of nucleic acids.
CC   The composition is used to deliver nucleic acids to eukaryotic cells, in
CC   vivo or in vitro, for expressing a therapeutic polypeptide or antisense
CC   molecule (or ribozyme). Specifically they are used for gene therapy of
CC   cancers (particularly non-small cell lung carcinoma), diabetes, cystic
CC   fibrosis and arteriosclerosis
XX
SQ   Sequence 51 AA;

  Query Match             84.0%;  Score 42;  DB 2;  Length 51;
  Best Local Similarity   90.0%;  Pred. No. 0.68;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db          6 TRLTRKRGLK 15


RESULT 14
ABB37687
ID   ABB37687 standard; peptide; 343 AA.
XX
AC   ABB37687;
XX
DT   04-FEB-2002  (first entry)
XX
DE   Peptide #5193 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 27; SEQ ID NO 30322; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a peptide encoded by a single exon nucleic acid probe
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 343 AA;

  Query Match             84.0%;  Score 42;  DB 4;  Length 343;
  Best Local Similarity   90.0%;  Pred. No. 5.1;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db        169 TRLTRKRGLK 178


RESULT 15
ABG52504
ID   ABG52504 standard; peptide; 343 AA.
XX
AC   ABG52504;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver peptide, SEQ ID No 31152.
XX
KW   Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW   hypercholesterolaemia; coronary heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
```

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 31152; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 343 AA;

Query Match 84.0%; Score 42; DB 4; Length 343;
Best Local Similarity 90.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTDKRGLK 10
              |||| |||||
Db        169 TRLTRKRGLK 178
```

Search completed: December 29, 2004, 12:28:50
Job time : 62.0227 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 9.65909 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-8
Perfect score: 50
Sequence: 1 TRLTDKRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	84.0	596	2	S32802	apolipoprotein B -
2	42	84.0	4563	1	LPHUB	apolipoprotein B-1
3	38	76.0	269	2	C60950	apolipoprotein B-1
4	38	76.0	779	2	JH0102	apolipoprotein B -
5	36	72.0	275	2	E60950	apolipoprotein B-1
6	35	70.0	170	2	A43654	probable periplasm
7	35	70.0	193	2	D90206	imidazoleglycerol-
8	35	70.0	354	2	AC2865	sugar binding prot
9	35	70.0	354	2	B97642	multiple sugar-bin
10	35	70.0	499	1	S17648	pyruvate kinase (E
11	35	70.0	499	2	S17649	pyruvate kinase (E
12	35	70.0	562	1	QQBEH5	phosphotransferase
13	35	70.0	563	2	T44214	probable phosphotr
14	35	70.0	563	2	T44029	ganciclovir kinase
15	35	70.0	612	2	B81246	glutamine-fructose
16	34	68.0	203	2	T44695	btuR protein [impo
17	34	68.0	258	2	T01873	hypothetical prote
18	34	68.0	612	2	H82022	glutamine-fructose
19	34	68.0	643	2	F97787	sodium/pantothenat
20	34	68.0	783	2	T00782	probable anthranil
21	34	68.0	1029	2	F86359	hypothetical prote
22	34	68.0	1073	2	T01955	hypothetical prote
23	34	68.0	1241	2	H84486	probable helicase
24	34	68.0	1265	2	F84517	probable helicase
25	34	68.0	1678	2	D86481	189.6K hypothetica
26	34	68.0	1752	2	T48965	hypothetical prote
27	33	66.0	235	2	G81138	probable succinate
28	33	66.0	360	2	S58205	DHR38 protein - si
29	33	66.0	406	2	AI0767	probable glycosylt
30	33	66.0	406	2	A90985	hypothetical prote
31	33	66.0	406	2	D85830	hypothetical prote
32	33	66.0	406	2	C64970	hypothetical prote
33	33	66.0	407	2	S52148	amsK protein - Erw
34	33	66.0	411	2	S15296	hypothetical prote
35	33	66.0	487	2	T21384	hypothetical prote
36	33	66.0	490	2	F82546	fimbrial assembly
37	33	66.0	1058	2	S65460	apolipoprotein B -
38	33	66.0	1252	2	A86501	RNA polymerase bet
39	33	66.0	1252	2	D72122	RNA polymerase bet
40	33	66.0	1252	2	G81686	DNA-directed RNA p
41	33	66.0	1252	2	H71529	DNA-directed RNA p
42	33	66.0	1262	2	F81548	DNA-directed RNA p
43	33	66.0	1366	2	T35985	probable large Pro
44	32	64.0	84	2	G87376	hypothetical prote
45	32	64.0	122	2	PN0152	DNA topoisomerase

ALIGNMENTS

RESULT 1
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchi
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional r
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g9301
C;Superfamily: apolipoprotein B

Query Match 84.0%; Score 42; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTDKRGLK 10
              |||| |||||
Db        226 TRLTRKRGLK 235
```

RESULT 2
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A2
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Sc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNI
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: r
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986

A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732

R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than o
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spe
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Ca
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place c
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

```
A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

  Query Match             84.0%;  Score 42;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 12;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db       3385 TRLTRKRGLK 3394


RESULT 3
C60950
apolipoprotein B-100 - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: C60950
A;Molecule type: DNA
A;Residues: 1-269 <LAW>
A;Cross-references: UNIPROT:Q60537; UNIPROT:Q60536
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

  Query Match             76.0%;  Score 38;  DB 2;  Length 269;
  Best Local Similarity   80.0%;  Pred. No. 5.5;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        216 SRLTRKRGLK 225


RESULT 4
JH0102
apolipoprotein B - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JH0102
R;Smith, T.J.
submitted to GenBank, June 1990
A;Reference number: A38864
A;Accession: JH0102
A;Molecule type: DNA
A;Residues: 1-779 <SMI>
A;Cross-references: UNIPROT:Q60536; GB:M35187
A;Note: this is a revision to the sequence from reference JH0101
R;Smith, T.J.; Hautamaa, D.; Maeda, N.
Gene 87, 309-310, 1990
A;Title: Sequence of the putative low-density lipoprotein receptor-binding regions of a
A;Reference number: JH0101; MUID:90236327; PMID:2332175
A;Contents: annotation
A;Note: this sequence has been revised in reference A38864
C;Genetics:
A;Gene: apoB
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein
F;435-445/Region: receptor binding
F;646-656/Region: receptor binding

  Query Match             76.0%;  Score 38;  DB 2;  Length 779;
  Best Local Similarity   80.0%;  Pred. No. 15;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        642 SRLTRKRGLK 651


RESULT 5
E60950
apolipoprotein B-100 - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: E60950
R;Law, A.; Scott, J.
J. Lipid Res. 31, 1109-1120, 1990
A;Title: A cross-species comparison of the apolipoprotein B domain that binds to the LD
A;Reference number: A60950; MUID:90324804; PMID:2373961
A;Accession: E60950
A;Molecule type: mRNA
A;Residues: 1-275 <LAW>
A;Cross-references: UNIPROT:Q7LZ77
C;Superfamily: apolipoprotein B
C;Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein

  Query Match             72.0%;  Score 36;  DB 2;  Length 275;
  Best Local Similarity   80.0%;  Pred. No. 15;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              | || |||||
Db        221 TSLTRKRGLK 230


RESULT 6
A43654
probable periplasmic receptor protein chvE precursor - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 09-Jul-2004
C;Accession: A43654
R;Huang, M.L.W.; Cangelosi, G.A.; Halperin, W.; Nester, E.W.
J. Bacteriol. 172, 1814-1822, 1990
A;Title: A chromosomal Agrobacterium tumefaciens gene required for effective plant sign
A;Reference number: A43654; MUID:90202696; PMID:2156804
A;Accession: A43654
A;Status: preliminary
```

```
A;Molecule type: DNA
A;Residues: 1-170 <HUA>
A;Cross-references: UNIPROT:P25548; GB:M30318

  Query Match             70.0%;  Score 35;  DB 2;  Length 170;
  Best Local Similarity   70.0%;  Pred. No. 15;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              | :||| |||
Db        141 TSITDKLGLK 150


RESULT 7
D90206
imidazoleglycerol-phosphate dehydratase (hisB) [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: D90206
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: D90206
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-193 <KUR>
A;Cross-references: UNIPROT:O33773; GB:AE006641; NID:g13813761; PIDN:AAK40907.1; GSPDB:G
C;Genetics:
A;Gene: hisB
C;Superfamily: imidazoleglycerol-phosphate dehydratase; imidazoleglycerol-phosphate dehy

  Query Match             70.0%;  Score 35;  DB 2;  Length 193;
  Best Local Similarity   70.0%;  Pred. No. 17;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              | | ||||:|
Db         81 TALGDKRGIK 90


RESULT 8
AC2865
sugar binding protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AC2865
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2865
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-354 <KUR>
A;Cross-references: UNIPROT:P25548; GB:AE008688; PIDN:AAL43337.1; PID:g17740831; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: chvE
A;Map position: circular chromosome

  Query Match             70.0%;  Score 35;  DB 2;  Length 354;
  Best Local Similarity   70.0%;  Pred. No. 30;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              | :||| |||
Db        141 TSITDKLGLK 150


RESULT 9
B97642
multiple sugar-binding periplasmic receptor chve precursor [imported] - Agrobacterium t
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97642
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B ;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97642
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-354 <KUR>
A;Cross-references: UNIPROT:P25548; GB:AE007869; PIDN:AAK88091.1; PID:g15157521; GSPDB:G
C;Genetics:
A;Gene: AGR_C_4267
A;Map position: circular chromosome

  Query Match             70.0%;  Score 35;  DB 2;  Length 354;
  Best Local Similarity   70.0%;  Pred. No. 30;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              | :||| |||
Db        141 TSITDKLGLK 150


RESULT 10
S17648
pyruvate kinase (EC 2.7.1.40) isoform 1 - Trypanosoma brucei
C;Species: Trypanosoma brucei
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S17648
R;Allert, S.; Ernest, I.; Poliszczak, A.; Opperdoes, F.R.; Michels, P.A.M.
Eur. J. Biochem. 200, 19-27, 1991
A;Title: Molecular cloning and analysis of two tandemly linked genes for pyruvate kinase
A;Reference number: S17648; MUID:91348039; PMID:1879424
A;Accession: S17648
A;Molecule type: DNA
A;Residues: 1-499 <ALL>
A;Cross-references: UNIPROT:P30615; EMBL:X57950; NID:g10947; PIDN:CAA41018.1; PID:g10948
A;Experimental source: strain 427
C;Function:
A;Description: catalyzes the transphosphorylation of phosphoenolpyruvate and ADP to pyru
A;Pathway: glycolysis
C;Superfamily: pyruvate kinase
C;Keywords: ATP biosynthesis; glycolysis; magnesium; metalloprotein; phosphotransferase;
F;50,212,263/Binding site: substrate phosphate (Arg, Ser, Arg) #status predicted
F;239/Active site: Lys #status predicted
F;298,333/Binding site: potassium (Gln, Glu) #status predicted

  Query Match             70.0%;  Score 35;  DB 1;  Length 499;
  Best Local Similarity   75.0%;  Pred. No. 41;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGL 9
              ||||:||:
Db        171 RLTDRRGI 178


RESULT 11
S17649
pyruvate kinase (EC 2.7.1.40) isoform 2 - Trypanosoma brucei
C;Species: Trypanosoma brucei
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S17649
R;Allert, S.; Ernest, I.; Poliszczak, A.; Opperdoes, F.R.; Michels, P.A.M.
Eur. J. Biochem. 200, 19-27, 1991
```

```
A;Title: Molecular cloning and analysis of two tandemly linked genes for pyruvate kinase
A;Reference number: S17648; MUID:91348039; PMID:1879424
A;Accession: S17649
A;Molecule type: DNA
A;Residues: 1-499 <ALL>
A;Cross-references: UNIPROT:P30616; EMBL:X57951; NID:g10949; PIDN:CAA41019.1; PID:g10950
A;Experimental source: strain 427
C;Function:
A;Description: catalyzes the transphosphorylation of phosphoenolpyruvate and ADP to pyru
A;Pathway: glycolysis
C;Superfamily: pyruvate kinase
C;Keywords: ATP biosynthesis; glycolysis; magnesium; metalloprotein; phosphotransferase;
F;50,212,263/Binding site: substrate phosphate (Arg, Ser, Arg) #status predicted
F;239/Active site: Lys #status predicted
F;298,333/Binding site: potassium (Gln, Glu) #status predicted

  Query Match             70.0%;  Score 35;  DB 2;  Length 499;
  Best Local Similarity   75.0%;  Pred. No. 41;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGL 9
              ||||:||:
Db        171 RLTDRRGI 178


RESULT 12
QQBEH5
phosphotransferase (EC 2.7.1.-) - human herpesvirus 6 (strain Uganda-1102)
C;Species: human herpesvirus 6
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: E36769
R;Lawrence, G.L.; Chee, M.; Craxton, M.A.; Gompels, U.A.; Honess, R.W.; Barrell, B.G.
J. Virol. 64, 287-299, 1990
A;Title: Human herpesvirus 6 is closely related to human cytomegalovirus.
A;Reference number: A33560; MUID:90080132; PMID:2152817
A;Accession: E36769
A;Molecule type: DNA
A;Residues: 1-562 <LAW>
A;Cross-references: UNIPROT:P24446; GB:M68963; GB:M28243; NID:g325494; PIDN:AAA65577.1;
C;Superfamily: human cytomegalovirus phosphotransferase
C;Keywords: phosphotransferase

  Query Match             70.0%;  Score 35;  DB 1;  Length 562;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGLK 10
              |:|||| :
Db        421 RMTDKRGCR 429


RESULT 13
T44214
probable phosphotransferase (EC 2.7.1.-) U69 [similarity] - human herpesvirus 6 (strain
C;Species: human herpesvirus 6
A;Variety: strain Z29
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 05-May-2000
C;Accession: T44214
R;Dominguez, G.; Dambaugh, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.
J. Virol. 73, 8040-8052, 1999
A;Title: Human herpesvirus 6B genome sequence: coding content and comparison with human
A;Reference number: Z22734; MUID:99412318; PMID:10482553
A;Accession: T44214
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-563 <DOM>
A;Cross-references: EMBL:AF157706; PIDN:AAD49670.1
A;Experimental source: strain Z29; variant B
C;Genetics:
A;Note: U69
C;Superfamily: human cytomegalovirus phosphotransferase
C;Keywords: phosphotransferase

  Query Match             70.0%;  Score 35;  DB 2;  Length 563;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGLK 10
              |:|||| :
Db        422 RMTDKRGCR 430


RESULT 14
T44029
ganciclovir kinase [imported] - human herpesvirus 6 (strain HST)
C;Species: human herpesvirus 6
A;Variety: strain HST
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44029
R;Isegawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kaw
J. Virol. 73, 8053-8063, 1999
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and
A;Reference number: Z22732; MUID:99412319; PMID:10482554
A;Accession: T44029
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-563 <ISE>
A;Cross-references: UNIPROT:Q9WT04; EMBL:AB021506; NID:g4995977; PIDN:BAA78290.1; PID:g
A;Experimental source: strain HST; pop. variant B
C;Genetics:
A;Note: U69
C;Superfamily: human cytomegalovirus phosphotransferase

  Query Match             70.0%;  Score 35;  DB 2;  Length 563;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGLK 10
              |:|||| :
Db        422 RMTDKRGCR 430


RESULT 15
B81246
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) NMB0031 [simila
N;Alternate names: glucosamine fructose-6-phosphate aminotransferase [misnomer]
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: B81246
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: B81246
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-612 <TET>
A;Cross-references: UNIPROT:Q9K1P9; GB:AE002361; GB:AE002098; NID:g7225245; PIDN:AAF405
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0031
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
F;2-612/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status pred
F;2/Active site: Cys #status predicted

  Query Match             70.0%;  Score 35;  DB 2;  Length 612;
  Best Local Similarity   87.5%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGL 9
              ||||| ||
Db        224 RLTDKNGL 231
```

Search completed: December 29, 2004, 12:39:07
Job time : 11.6591 secs

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 58.4091 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-8
Perfect score: 50
Sequence: 1 TRLTDKRGLK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	84.0	414	2	Q7YQR5	Q7yqr5 aotus vocif
2	42	84.0	596	2	Q28473	Q28473 macaca fasc
3	42	84.0	3262	2	Q13788	Q13788 homo sapien
4	42	84.0	4563	1	APB_HUMAN	P04114 homo sapien
5	42	84.0	4563	2	Q7Z600	Q7z600 homo sapien
6	39	78.0	132	2	Q8LCX0	Q8lcx0 arabidopsis
7	38	76.0	421	2	Q7TN68	Q7tn68 glaucomys v
8	38	76.0	430	2	Q7NHS0	Q7nhs0 gloeobacter
9	38	76.0	432	2	Q7YR10	Q7yr10 diceros bic
10	38	76.0	436	2	Q7YQM8	Q7yqm8 nyctimene a
11	38	76.0	438	2	Q7YQM7	Q7yqm7 pteropus hy
12	38	76.0	438	2	Q7YR04	Q7yr04 rousettus a
13	38	76.0	445	2	Q7YR08	Q7yr08 chaetophrac
14	38	76.0	445	2	Q7TN64	Q7tn64 agouti paca
15	38	76.0	445	2	Q7TN71	Q7tn71 hydrochoeru
16	38	76.0	445	2	Q7TN72	Q7tn72 erethizon d
17	38	76.0	780	2	Q60536	Q60536 mesocricetu
18	38	76.0	780	2	Q60537	Q60537 mesocricetu
19	36	72.0	231	2	Q6N7G3	Q6n7g3 rhodopseudo
20	36	72.0	231	2	CAE27735	Cae27735 rhodopseu
21	36	72.0	275	2	Q7LZ77	Q7lz77 gallus gall
22	36	72.0	387	2	Q7YQN2	Q7yqn2 phalanger o
23	36	72.0	400	2	Q7YQM9	Q7yqm9 ornithorhyn
24	36	72.0	405	2	Q7YQN0	Q7yqn0 tachyglossu
25	36	72.0	445	2	Q7TN70	Q7tn70 dinomys bra
26	36	72.0	451	2	Q9KXP4	Q9kxp4 streptomyce
27	36	72.0	453	2	Q827S8	Q827s8 streptomyce
28	36	72.0	995	2	Q7R285	Q7r285 giardia lam
29	35	70.0	133	2	Q6U2N6	Q6u2n6 infectious
30	35	70.0	133	2	AAQ75361	Aaq75361 infectiou
31	35	70.0	182	2	Q6L6Z2	Q6l6z2 thermoprote
32	35	70.0	182	2	BAD18908	Bad18908 thermopro
33	35	70.0	192	2	Q6L705	Q6l705 thermoprote
34	35	70.0	192	2	BAD18895	Bad18895 thermopro
35	35	70.0	193	1	HIS7_SULSO	O33773 sulfolobus
36	35	70.0	225	2	Q8A5R4	Q8a5r4 bacteroides
37	35	70.0	265	2	Q72JP7	Q72jp7 thermus the
38	35	70.0	265	2	AAS81069	Aas81069 thermus t
39	35	70.0	278	2	Q9CQM3	Q9cqm3 m mus muscu
40	35	70.0	278	2	Q91WA1	Q91wa1 mus musculu
41	35	70.0	354	1	CHVE_AGRT5	P25548 agrobacteri
42	35	70.0	407	2	Q7TN65	Q7tn65 atherurus a
43	35	70.0	412	2	Q7TN69	Q7tn69 hystrix bra
44	35	70.0	499	1	KPY1_TRYBB	P30615 trypanosoma
45	35	70.0	499	1	KPY2_TRYBB	P30616 trypanosoma

ALIGNMENTS

RESULT 1
Q7YQR5
ID Q7YQR5 PRELIMINARY; PRT; 414 AA.
AC Q7YQR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein B 100 (Fragment).
GN Name=apoB-100;
OS Aotus vociferans (Spix's owl monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=57176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22761261; PubMed=12878460;
RA Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT "A new phylogenetic marker, apolipoprotein B, provides compelling
RT evidence for eutherian relationships.";
RL Mol. Phylogenet. Evol. 28:225-240(2003).
DR EMBL; AF548396; AAP97352.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 45955 MW; EEFA8492157E1BDE CRC64;

Query Match 84.0%; Score 42; DB 2; Length 414;
Best Local Similarity 90.0%; Pred. No. 7.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTDKRGLK 10
              |||| |||||
Db        258 TRLTRKRGLK 267
```

RESULT 2
Q28473
ID Q28473 PRELIMINARY; PRT; 596 AA.
AC Q28473;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Apolipoprotein B (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92075708; PubMed=1742325;
RA Pape M.E., Castle C.K., Murray R.W., Funk G.M., Hunt C.E.,
RA Marotti K.R., Melchior G.W.;

```
RT   "Apo B metabolism in the cynomolgus monkey: evidence for post-
RT   transcriptional regulation.";
RL   Biochim. Biophys. Acta 1086:326-334(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Murray R.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X15737; CAA33755.1; -.
DR   PIR; S32802; S32802.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     596    596
SQ   SEQUENCE   596 AA;  66757 MW;  B13BBA74E25C3120 CRC64;

  Query Match             84.0%;  Score 42;  DB 2;  Length 596;
  Best Local Similarity   90.0%;  Pred. No. 11;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db        226 TRLTRKRGLK 235


RESULT 3
Q13788
ID   Q13788      PRELIMINARY;      PRT;  3262 AA.
AC   Q13788;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   APOB protein (Fragment).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191999; PubMed=2883086;
RA   Carlsson P., Darnfors C., Olofsson S.O., Bjursell G.;
RT   "Analysis of the human apolipoprotein B gene; complete structure of
RT   the B-74 region.";
RL   Gene 49:29-51(1986).
DR   EMBL; M15421; AAA51758.1; -.
DR   PIR; A27850; LPHUB.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005319; F:lipid transporter activity; NAS.
DR   GO; GO:0006869; P:lipid transport; NAS.
FT   NON_TER       1      1
SQ   SEQUENCE   3262 AA;  370140 MW;  56603BC0618DD40D CRC64;

  Query Match             84.0%;  Score 42;  DB 2;  Length 3262;
  Best Local Similarity   90.0%;  Pred. No. 68;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db       2084 TRLTRKRGLK 2093


RESULT 4
APB_HUMAN
ID   APB_HUMAN      STANDARD;      PRT;  4563 AA.
AC   P04114; O00502; Q13787;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein
DE   B-48 (Apo B-48)].
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87016385; PubMed=3763409;
RA   Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.,
RA   Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
RT   "Complete cDNA and derived protein sequence of human apolipoprotein B-
RT   100.";
RL   Nucleic Acids Res. 14:7501-7503(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=88003974; PubMed=3652907;
RA   Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
RA   Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
RT   "DNA sequence of the human apolipoprotein B gene.";
RL   DNA 6:363-372(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS ILE-98 AND GLU-4181.
RX   MEDLINE=87008488; PubMed=3759943;
RA   Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
RA   Gotto A.M. Jr., Chan L.;
RT   "The complete cDNA and amino acid sequence of human apolipoprotein B-
RT   100.";
RL   J. Biol. Chem. 261:12918-12921(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87041416; PubMed=3464946;
RA   Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
RA   Lee N., Brewer H.B. Jr.;
RT   "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
RT   derived amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87161758; PubMed=3030729;
RA   Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
RA   Zannis V.I.;
RT   "The complete sequence and structural analysis of human apolipoprotein
RT   B-100: relationship between apoB-100 and apoB-48 forms.";
RL   EMBO J. 5:3495-3507(1986).
RN   [6]
RP   SEQUENCE OF 709-906 FROM N.A.
RX   MEDLINE=85270450; PubMed=3860836;
RA   Deeb S.S., Motulsky A.G., Albers J.J.;
RT   "A partial cDNA clone for human apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
RN   [7]
RP   SEQUENCE OF 3056-3159 FROM N.A.
RX   MEDLINE=86041888; PubMed=3903660;
RA   Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
RA   Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
RT   "Human apolipoprotein B: identification of cDNA clones and
RT   characterization of mRNA.";
RL   Nucleic Acids Res. 13:6937-6953(1985).
RN   [8]
RP   SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A., AND VARIANT GLU-4181.
RX   MEDLINE=86093680; PubMed=3841204;
RA   Carlsson P., Olofsson S.O., Bondjers G., Darnfors C., Wiklund O.,
RA   Bjursell G.;
RT   "Molecular cloning of human apolipoprotein B cDNA.";
RL   Nucleic Acids Res. 13:8813-8826(1985).
RN   [9]
RP   SEQUENCE OF 3109-4563 FROM N.A.
RX   MEDLINE=85300528; PubMed=2994225;
RA   Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F., Urdea M.S.,
RA   Levy-Wilson B., Powell L.M., Pease R.J., Eddy R., Nakai H., Byers M.,
RA   Priestley L.M., Robertson E., Rall L.B., Betsholtz C., Shows T.B.,
RA   Mahley R.W., Scott J.;
RT   "Human apolipoprotein B: structure of carboxyl-terminal domains, sites
RT   of gene expression, and chromosomal localization.";
RL   Science 230:37-43(1985).
RN   [10]
```

```
RP   SEQUENCE OF 1-291 FROM N.A.
RX   MEDLINE=86149325; PubMed=3513177;
RA   Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
RA   Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
RT   "Isolation of a cDNA clone encoding the amino-terminal region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
RN   [11]
RP   SEQUENCE OF 1-1670 FROM N.A., AND VARIANT ILE-98.
RX   MEDLINE=86287319; PubMed=3461454;
RA   Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W., Yamanaka M.,
RA   Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
RT   "Analysis of cDNA clones encoding the entire B-26 region of human
RT   apolipoprotein B.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
RN   [12]
RP   PARTIAL SEQUENCE, AND IDENTIFICATION OF APO-B48.
RX   MEDLINE=88018019; PubMed=3659919;
RA   Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
RA   Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
RA   Gotto A.M. Jr., Li W.-H., Chan L.;
RT   "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
RT   specific in-frame stop codon.";
RL   Science 238:363-366(1987).
RN   [13]
RP   DOMAINS.
RX   MEDLINE=87039351; PubMed=3773997;
RA   Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
RA   Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
RA   Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
RA   Levy-Wilson B., Scott J.;
RT   "Complete protein sequence and identification of structural domains of
RT   human apolipoprotein B.";
RL   Nature 323:734-738(1986).
RN   [14]
RP   DOMAINS.
RA   Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
RA   Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
RA   Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
RT   "Sequence, structure, receptor-binding domains and internal repeats of
RT   human apolipoprotein B-100.";
RL   Nature 323:738-742(1986).
RN   [15]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=86242245; PubMed=3087360;
RA   Dashti N., Lee D.M., Mok T.;
RT   "Apolipoprotein B is a calcium binding protein.";
RL   Biochem. Biophys. Res. Commun. 137:493-499(1986).
RN   [16]
RP   PALMITOYLATION OF CYS-1112.
RX   MEDLINE=20143590; PubMed=10679026;
RA   Zhao Y., McCabe J.B., Vance J., Berthiaume L.G.;
RT   "Palmitoylation of apolipoprotein B is required for proper
RT   intracellular sorting and transport of cholesteroyl esters and
RT   triglycerides.";
RL   Mol. Biol. Cell 11:721-734(2000).
RN   [17]
RP   VARIANT SER-4338.
RX   MEDLINE=91071750; PubMed=1979313;
RA   Navajas M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
RA   Cunny G., Cambien F., Roizes G.;
RT   "Detection by denaturing gradient gel electrophoresis of a new
RT   polymorphism in the apolipoprotein B gene.";
RL   Hum. Genet. 86:91-93(1990).
RN   [18]
RP   VARIANT FDB GLN-3527.
RX   MEDLINE=89098975; PubMed=2563166;
RA   Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
RA   McCarthy B.J.;
RT   "Association between a specific apolipoprotein B mutation and familial
RT   defective apolipoprotein B-100.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
RN   [19]
RP   VARIANT LEU-2739.
RX   MEDLINE=91016974; PubMed=2216805;
RA   Huang L.-S., Gavish D., Breslow J.L.;
RT   "Sequence polymorphism in the human apoB gene at position 8344.";
RL   Nucleic Acids Res. 18:5922-5922(1990).
RN   [20]
RP   VARIANT FDB CYS-3558.
RX   MEDLINE=95190020; PubMed=7883971;
RA   Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
RA   Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
RT   "Familial ligand-defective apolipoprotein B. Identification of a new
RT   mutation that decreases LDL receptor binding affinity.";
RL   J. Clin. Invest. 95:1225-1234(1995).
RN   [21]
RP   VARIANTS LEU-1437; SER-1914; LYS-2566; THR-3121; ALA-3945; MET-4128
RP   AND THR-4481.
RX   MEDLINE=97044521; PubMed=8889592;
RA   Poirier O., Ricard S., Behague I., Souriau C., Evans A.E.,
RA   Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
RT   "Detection of new variants in the apolipoprotein B (Apo B) gene by
RT   PCR-SSCP.";
RL   Hum. Mutat. 8:282-285(1996).
RN   [22]
RP   VARIANTS FDB GLN-3527 AND CYS-3558.
RX   MEDLINE=97403938; PubMed=9259199;
RA   Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
RA   Krempf M., Giraudet P., Junien C., Boileau C.;
RT   "Familial ligand-defective apolipoprotein B-100: simultaneous
RT   detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French
RT   population.";
RL   Hum. Mutat. 10:160-163(1997).
RN   [23]
RP   VARIANTS SER-1914; ARG-1923; LEU-2739; ASP-3319; THR-3427; GLN-3432
RP   AND ILE-3921.
RX   MEDLINE=98141125; PubMed=9490296;
RA   Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
RT   "Screening for mutations of the apolipoprotein B gene causing
RT   hypocholesterolemia.";
RL   Hum. Genet. 102:44-49(1998).
CC   -!- FUNCTION: Apolipoprotein B is a major protein constituent of
CC       chylomicrons (apo B-48), LDL (apo B-100) and VLDL (apo B-100). Apo
CC       B-100 functions as a recognition signal for the cellular binding
CC       and internalization of LDL particles by the apoB/E receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.

  Query Match             84.0%;  Score 42;  DB 1;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 97;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db       3385 TRLTRKRGLK 3394


RESULT 5
Q7Z600
ID   Q7Z600      PRELIMINARY;      PRT;  4563 AA.
AC   Q7Z600;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Including Ag(X) antigen).
GN   Name=APOB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY324608; AAP72970.1; -.
DR   GO; GO:0005319; F:lipid transporter activity; IEA.
DR   GO; GO:0006869; P:lipid transport; IEA.
DR   InterPro; IPR009454; DUF1081.
DR   InterPro; IPR001747; Lipid_transprt_N.
DR   Pfam; PF06448; DUF1081; 1.
DR   Pfam; PF01347; Vitellogenin_N; 1.
DR   SMART; SM00638; LPD_N; 1.
KW   Lipoprotein.
SQ   SEQUENCE   4563 AA;  515553 MW;  030B34167CEDC63C CRC64;

  Query Match          84.0%;  Score 42;  DB 2;  Length 4563;
  Best Local Similarity   90.0%;  Pred. No. 97;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              |||| |||||
Db       3385 TRLTRKRGLK 3394


RESULT 6
Q8LCX0
ID   Q8LCX0      PRELIMINARY;      PRT;   132 AA.
AC   Q8LCX0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22088475; PubMed=12093376;
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY086357; AAM64425.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   132 AA;  14674 MW;  A3698270AE88CD31 CRC64;

  Query Match          78.0%;  Score 39;  DB 2;  Length 132;
  Best Local Similarity   70.0%;  Pred. No. 9.7;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              ||: ||||:|
Db         41 TRIIDKRGVK 50


RESULT 7
Q7TN68
ID   Q7TN68      PRELIMINARY;      PRT;   421 AA.
AC   Q7TN68;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Glaucomys volans (Southern flying squirrel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Petauristinae;
OC   Glaucomys.
OX   NCBI_TaxID=64683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243379; AAP50767.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     421    421
SQ   SEQUENCE   421 AA;  46747 MW;  D47B77BD4F864FD1 CRC64;

  Query Match          76.0%;  Score 38;  DB 2;  Length 421;
  Best Local Similarity   80.0%;  Pred. No. 52;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        264 SRLTRKRGLK 273


RESULT 8
Q7NHS0
ID   Q7NHS0      PRELIMINARY;      PRT;   430 AA.
AC   Q7NHS0;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Gll2465 protein.
GN   OrderedLocusNames=gll2465;
OS   Gloeobacter violaceus.
OC   Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OX   NCBI_TaxID=33072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PCC 7421;
RX   MEDLINE=22977040; PubMed=14621292;
RA   Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA   Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA   Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT   cyanobacterium that lacks thylakoids.";
RL   DNA Res. 10:137-145(2003).
DR   EMBL; AP006576; BAC90406.1; -.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR001296; Glyco_trans_1.
DR   Pfam; PF00534; Glycos_transf_1; 1.
KW   Complete proteome.
SQ   SEQUENCE   430 AA;  47641 MW;  0F8D1A8F7A38342D CRC64;

  Query Match          76.0%;  Score 38;  DB 2;  Length 430;
  Best Local Similarity   77.8%;  Pred. No. 53;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTDKRGLK 10
              |||||:||:
Db        231 RLTDKKGLE 239


RESULT 9
Q7YR10
ID   Q7YR10      PRELIMINARY;      PRT;   432 AA.
AC   Q7YR10;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243375; AAP50763.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     432    432
SQ   SEQUENCE   432 AA;  48171 MW;  F27B7AB39604732C CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 432;
  Best Local Similarity   80.0%;  Pred. No. 54;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        275 SRLTRKRGLK 284


RESULT 10
Q7YQM8
ID   Q7YQM8      PRELIMINARY;      PRT;   436 AA.
AC   Q7YQM8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Nyctimene albiventer (Common tube-nosed fruit bat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Nyctimene.
OX   NCBI_TaxID=48988;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548435; AAP97391.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     436    436
SQ   SEQUENCE   436 AA;  48717 MW;  1C4A7EAD72D2C629 CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 436;
  Best Local Similarity   80.0%;  Pred. No. 54;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        279 SRLTRKRGLK 288


RESULT 11
Q7YQM7
ID   Q7YQM7      PRELIMINARY;      PRT;   438 AA.
AC   Q7YQM7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Pteropus hypomelanus (Small flying fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Pteropus.
OX   NCBI_TaxID=9405;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548436; AAP97392.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48734 MW;  2BD85BCBF4E2CC41 CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 54;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        281 SRLTRKRGLK 290


RESULT 12
Q7YR04
ID   Q7YR04      PRELIMINARY;      PRT;   438 AA.
AC   Q7YR04;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Rousettus amplexicaudatus (Common rousette).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
OC   Pteropodinae; Rousettus.
OX   NCBI_TaxID=58083;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243383; AAP50771.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     438    438
SQ   SEQUENCE   438 AA;  48597 MW;  41C890DEAF95C872 CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 54;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        281 SRLTRKRGLK 290


RESULT 13
Q7YR08
ID   Q7YR08      PRELIMINARY;      PRT;   445 AA.
AC   Q7YR08;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Chaetophractus villosus (South American armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Chaetophractus.
OX   NCBI_TaxID=29080;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
```

```
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243378; AAP50766.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49564 MW;  2DA5DC3ED2F0FDD2 CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 55;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


RESULT 14
Q7TN64
ID   Q7TN64      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN64;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Apolipoprotein B 100 (Fragment).
GN   Name=apoB-100;
OS   Agouti paca (Paca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Agoutidae; Agouti.
OX   NCBI_TaxID=108852;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AF548417; AAP97373.1; -.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49721 MW;  34AF7ABE90F121EF CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 55;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


RESULT 15
Q7TN71
ID   Q7TN71      PRELIMINARY;      PRT;   445 AA.
AC   Q7TN71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Apolipoprotein B (Fragment).
OS   Hydrochoerus hydrochaeris (Capybara) (Carpincho).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Hydrochaeridae;
OC   Hydrochaeris.
OX   NCBI_TaxID=10149;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22761261; PubMed=12878460;
RA   Amrine-Madsen H., Koepfli K.-P., Wayne R.K., Springer M.S.;
RT   "A new phylogenetic marker, apolipoprotein B, provides compelling
RT   evidence for eutherian relationships.";
RL   Mol. Phylogenet. Evol. 28:225-240(2003).
DR   EMBL; AY243369; AAP50757.1; -.
DR   InterPro; IPR000871; Beta_lactamase_A.
DR   PROSITE; PS00146; BETA_LACTAMASE_A; UNKNOWN_1.
KW   Lipoprotein.
FT   NON_TER       1      1
FT   NON_TER     445    445
SQ   SEQUENCE   445 AA;  49520 MW;  CBBA2DD53D7A18D2 CRC64;

  Query Match             76.0%;  Score 38;  DB 2;  Length 445;
  Best Local Similarity   80.0%;  Pred. No. 55;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTDKRGLK 10
              :||| |||||
Db        288 SRLTRKRGLK 297


Search completed: December 29, 2004, 12:37:36
Job time : 60.5202 secs
```

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 54.9205 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-13
Perfect score: 44
Sequence: 1 TRLTRRGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	9	2	AAY30694	Aay30694 Apo-B100
2	41	93.2	9	2	AAY30695	Aay30695 Apo-B100
3	41	93.2	9	2	AAY30696	Aay30696 Apo-B100
4	35	79.5	404	6	ABR43240	Abr43240 Human PMM
5	35	79.5	548	5	ABG97506	Abg97506 Human NOV
6	34	77.3	151	8	ABO58868	Abo58868 Human gen
7	34	77.3	311	8	ADK71826	Adk71826 Human kin
8	34	77.3	933	8	ADP29449	Adp29449 Human sec
9	34	77.3	984	7	ADE38441	Ade38441 Human pro
10	34	77.3	984	8	ADJ75552	Adj75552 Marker ge
11	33.5	76.1	10	2	AAY30684	Aay30684 Apo-B100
12	33.5	76.1	10	2	AAY30683	Aay30683 Apo-B100
13	33.5	76.1	10	2	AAY30686	Aay30686 Apo-B100
14	33.5	76.1	10	2	AAY30682	Aay30682 Apo-B100
15	33.5	76.1	10	2	AAY30685	Aay30685 Apo-B100
16	33.5	76.1	10	2	AAY30687	Aay30687 Apo-B100
17	33.5	76.1	11	2	AAW57205	Aaw57205 Apo B bin
18	33.5	76.1	13	2	AAW57207	Aaw57207 Apo B 100
19	33.5	76.1	15	2	AAW41261	Aaw41261 Apolipopr
20	33.5	76.1	15	2	AAW96892	Aaw96892 ApoB-100
21	33.5	76.1	20	6	ABJ37575	Abj37575 Heparin b
22	33.5	76.1	22	2	AAW57208	Aaw57208 Apo B 100
23	33.5	76.1	22	2	AAW57209	Aaw57209 Apo B 100
24	33.5	76.1	34	5	AAE14541	Aae14541 Human apo
25	33.5	76.1	36	2	AAW96876	Aaw96876 Nucleic a
26	33.5	76.1	37	2	AAW64587	Aaw64587 Human apo
27	33.5	76.1	51	2	AAW96845	Aaw96845 Nucleic a
28	33.5	76.1	343	4	ABB37687	Abb37687 Peptide #
29	33.5	76.1	343	4	ABG52504	Abg52504 Human liv
30	33.5	76.1	377	2	AAR72704	Aar72704 Human apo
31	33.5	76.1	377	2	AAR34031	Aar34031 Sequence
32	33.5	76.1	2463	8	ADJ57400	Adj57400 Human apo
33	33.5	76.1	3923	2	AAY31237	Aay31237 Human Apo
34	33.5	76.1	4536	2	AAW41262	Aaw41262 Apolipopr
35	33.5	76.1	4536	2	AAW96826	Aaw96826 Amino aci
36	33.5	76.1	4560	5	AAU98981	Aau98981 Human apo
37	33.5	76.1	4561	7	ADD48677	Add48677 Human Pro
38	33.5	76.1	4563	5	AAO15893	Aao15893 Human apo
39	33.5	76.1	4563	6	ABR40253	Abr40253 Human ali
40	33.5	76.1	4563	6	ABU79140	Abu79140 Apolipopr
41	33.5	76.1	4563	7	ADF43408	Adf43408 Apolipopr
42	33.5	76.1	4563	8	ADH18871	Adh18871 Human apo
43	33.5	76.1	4563	8	ADH18870	Adh18870 Human apo
44	33.5	76.1	4563	8	ADO33445	Ado33445 Human apo
45	33.5	76.1	4563	8	ADO33447	Ado33447 Human apo

ALIGNMENTS

RESULT 1
AAY30694
ID AAY30694 standard; peptide; 9 AA.
XX
AC AAY30694;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRRGLK 9
|||||||||
Db 1 TRLTRRGLK 9

RESULT 2
AAY30695
ID AAY30695 standard; peptide; 9 AA.
XX
AC AAY30695;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 9 AA;

Query Match 93.2%; Score 41; DB 2; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRRGLK 9
||||:||||
Db 1 TRLTKRGLK 9

RESULT 3
AAY30696
ID AAY30696 standard; peptide; 9 AA.
XX
AC AAY30696;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 9 AA;

Query Match 93.2%; Score 41; DB 2; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRLTRRGLK 9

```
                |||||:|||
Db          1 TRLTRKGLK 9


RESULT 4
ABR43240
ID   ABR43240 standard; protein; 404 AA.
XX
AC   ABR43240;
XX
DT   07-JUL-2003  (first entry)
XX
DE   Human PMMM-1 protein SEQ ID NO:1.
XX
KW   Human; protein modification and maintenance molecule; PMMM; cytostatic;
KW   antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective; AIDS;
KW   cerebroprotective; anti-HIV; antiallergic; antiinflammatory; cancer;
KW   thyromimetic; gene therapy; cell proliferative disorder; atherosclerosis;
KW   neurological disorder; epilepsy; Huntington's disease; stroke; allergy;
KW   immune disorder; inflammatory disorder; developmental disorder;
KW   hypothyroidism; Cushing's syndrome; infection.
XX
OS   Homo sapiens.
XX
PN   WO2003025131-A2.
XX
PD   27-MAR-2003.
XX
PF   13-SEP-2002; 2002WO-US029221.
XX
PR   14-SEP-2001; 2001US-0322196P.
PR   21-SEP-2001; 2001US-0324134P.
PR   05-OCT-2001; 2001US-0327233P.
PR   26-OCT-2001; 2001US-0346198P.
PR   02-NOV-2001; 2001US-0343980P.
PR   09-NOV-2001; 2001US-0348887P.
PR   16-NOV-2001; 2001US-0332423P.
PR   28-NOV-2001; 2001US-0334145P.
PR   28-NOV-2001; 2001US-0334229P.
PR   06-DEC-2001; 2001US-0337451P.
PR   25-JAN-2002; 2002US-0351928P.
PR   21-MAR-2002; 2002US-0366837P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sprague WW,  Chawla NK,  Warren BA,  Tang YT,  Elliott VS;
PI   Marquis JP,  Li JX,  Griffin JA,  Gietzen KJ,  Yang J,  Lu DAM;
PI   Emerling BM,  Duggan BM,  Richardson TW,  Lee SY,  Ramkumar J,  Becha SD;
PI   Lehr-Mason PM,  Swarnakar A,  Tran UK,  Kable AE,  Hafalia AJA,  Khare R;
XX
DR   WPI; 2003-354597/33.
DR   N-PSDB; ACC59959.
XX
PT   New human protein modification and maintenance molecules (PMMM), useful
PT   for diagnosing, treating and preventing diseases or conditions associated
PT   with the aberrant PMMM expression e.g. cancer, AIDS, epilepsy, or
PT   infections.
XX
PS   Claim 1; Page 206-207; 270pp; English.
XX
CC   ACC59959 to ACC59989 encode the human protein modification and
CC   maintenance molecule proteins given in ABR43240 to ABR43270, designated
CC   PMMM-1 to PMMM-31 (I). (I) have cytostatic, antiarteriosclerotic,
CC   anticonvulsant, nootropic, neuroprotective, cerebroprotective, anti-HIV,
CC   antiallergic, antiinflammatory and thyromimetic activities, and can be
CC   used in gene therapy. The PMMM polypeptides and polynucleotides are
CC   useful in diagnosing, treating and preventing diseases or conditions
CC   associated with the decreased expression or overexpression of PMMM, such
CC   as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC   epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC   allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)
CC   disorders, or infections. They are also useful in assessing the effects
CC   of exogenous compounds on the expression of nucleic acid and amino acid
CC   sequences of PMMM. The PMMMs or their fragments are useful in screening
CC   compounds for effectiveness as agonist or antagonist of the polypeptides,
CC   or in altering the expression of the target polynucleotide and compounds
CC   that specifically bind to or modulate the activity of the polypeptide.
CC   The microarray is useful in monitoring or measuring protein-protein
CC   interactions, drug-target interactions, and gene expression profiles
XX
SQ   Sequence 404 AA;

  Query Match             79.5%;  Score 35;  DB 6;  Length 404;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||:|:|||:
Db        367 TRITKRGLE 375


RESULT 5
ABG97506
ID   ABG97506 standard; protein; 548 AA.
XX
AC   ABG97506;
XX
DT   16-DEC-2002  (first entry)
XX
DE   Human NOVX25 protein.
XX
KW   Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
KW   Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
KW   tuberous sclerosis; hypercalcaemia; Parkinson's disease; depression;
KW   Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
KW   multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
KW   obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
KW   infertility; inflammatory bowel disease; atherosclerosis; hypertension;
KW   scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
KW   asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
KW   bacterial infection; parasitic infection; graft-versus-host disease;
KW   cell differentiation; cell proliferation; haematopoiesis; wound healing;
KW   angiogenesis.
XX
OS   Homo sapiens.
XX
PN   WO200272770-A2.
XX
PD   19-SEP-2002.
XX
PF   08-MAR-2002; 2002WO-US007283.
XX
PR   08-MAR-2001; 2001US-0274281P.
PR   09-MAR-2001; 2001US-0274849P.
PR   12-MAR-2001; 2001US-0275235P.
PR   13-MAR-2001; 2001US-0275579P.
PR   13-MAR-2001; 2001US-0275601P.
PR   14-MAR-2001; 2001US-0276000P.
PR   20-MAR-2001; 2001US-0277239P.
PR   20-MAR-2001; 2001US-0277327P.
PR   20-MAR-2001; 2001US-0277338P.
PR   21-MAR-2001; 2001US-0277791P.
PR   22-MAR-2001; 2001US-0277833P.
PR   23-MAR-2001; 2001US-0278152P.
PR   26-MAR-2001; 2001US-0278894P.
PR   27-MAR-2001; 2001US-0279036P.
PR   28-MAR-2001; 2001US-0279344P.
PR   30-MAR-2001; 2001US-0280233P.
PR   02-APR-2001; 2001US-0280802P.
PR   02-MAY-2001; 2001US-0288148P.
PR   31-MAY-2001; 2001US-0294821P.
PR   31-OCT-2001; 2001US-0335302P.
PR   04-DEC-2001; 2001US-0338375P.
PR   07-MAR-2002; 2002US-00094466.
XX
PA   (CURA-) CURAGEN CORP.
```

```
XX
PI   Spytek KA,  Vernet CA,  Tchernev VT,  Malyankar UM,  Gerlach VL;
PI   Li L,  Zerhusen BD,  Patturajan M,  Gusev VY,  Kekuda R,  Pena CEA;
PI   Zhong M,  Gangolli EA,  Taupier RJ;
XX
DR   WPI; 2002-713508/77.
DR   N-PSDB; ABS78750.
XX
PT   New NOVX polypeptides and polynucleotides, useful for preventing,
PT   diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
PT   sclerosis, atherosclerosis, cancer, infections, osteoporosis or
PT   Parkinson's disease.
XX
PS   Claim 1; Page 161; 266pp; English.
XX
CC   The present invention relates to a new polypeptide (NOVX). The NOVX
CC   polypeptide, nucleic acid and antibody are useful in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease,
CC   preferably a NOVX-associated disorder. The NOVX nucleic acids,
CC   polypeptides and antibodies are useful for treating, preventing or
CC   diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
CC   syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC   hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC   palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
CC   telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC   obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC   infertility, inflammatory bowel disease, atherosclerosis, hypertension,
CC   scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
CC   asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
CC   infections, or graft-versus-host disease. The nucleic acids and
CC   polypeptides may also be used as targets for the identification of small
CC   molecules that modulate or inhibit e.g. neurogenesis, cell
CC   differentiation, cell proliferation, haematopoiesis, wound healing and
CC   angiogenesis, in gene therapy, in generation of antibodies that bind
CC   immunospecifically to NOVX substances for use in therapeutic or
CC   diagnostic methods. The nucleic acids are further used as hybridisation
CC   probes, in chromosome mapping, tissue typing, preventive medicine, and
CC   pharmacogenomics. The present amino acid sequence represents a human NOVX
CC   protein of the invention
XX
SQ   Sequence 548 AA;

  Query Match             79.5%;  Score 35;  DB 5;  Length 548;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||:|:|||:
Db        515 TRITKRGLE 523


RESULT 6
ABO58868
ID   ABO58868 standard; protein; 151 AA.
XX
AC   ABO58868;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon protein #5102.
XX
KW   Human; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 45; SEQ ID NO 32502; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe protein of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 151 AA;

  Query Match             77.3%;  Score 34;  DB 8;  Length 151;
  Best Local Similarity   100.0%;  Pred. No. 65;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTRRGLK 9
              |||||||
Db        119 LTRRGLK 125


RESULT 7
ADK71826
ID   ADK71826 standard; protein; 311 AA.
XX
AC   ADK71826;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human kinase and phosphatase KPP-3 protein.
XX
```

KW human; kinase; phosphatase; KPP; cardiovascular; antiarteriosclerotic;
KW hypotensive; vasotropic; antiinflammatory; antianginal; anti-HIV;
KW antiallergic; antiasthmatic; immunosuppressive; antithyroid;
KW dermatological; antidiabetic; nephrotropic; antigout; gastrointestinal;
KW neuroprotective; osteopathic; antiarthritic; uropathic; ophthalmological;
KW antirheumatic; antiparkinsonian; nootropic; anticonvulsant; hepatotropic;
KW antipsoriatic; haemostatic; cytostatic; antilipaemic; antiparasitic;
KW antihelmintic; antibacterial; virucide; protozoacide; fungicide;
KW cardiovascular disease; immune system; neurological; growth; development;
KW cell proliferation; viral; bacterial; fungal; parasitic; protozoan;
KW helminthic infection; transgenic; gene therapy; enzyme;
KW single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
PN WO2004018641-A2.
XX
PD 04-MAR-2004.
XX
PF 25-AUG-2003; 2003WO-US026635.
XX
PR 26-AUG-2002; 2002US-0406172P.
PR 25-SEP-2002; 2002US-0413910P.
PR 27-SEP-2002; 2002US-0414296P.
PR 11-OCT-2002; 2002US-0417821P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Baughn MR, Richardson TW, Marquis JP, Swarnakar A, Tang YT;
PI Becha SD, Emerling BM, Jin P, Wilson AD, Yue H, Gietzen KJ;
PI Chang H, Yang YG, Lee SY, Khare R, Elliott VS, Hafalia AJA;
PI Chawla NK, Ramkumar J, Gururajan R, Tribouley CM, Chien D, Tran UK;
PI Murage J;
XX
DR WPI; 2004-226830/21.
DR N-PSDB; ADK71885.
XX
PT New human kinases and phosphatases, useful for diagnosing, treating or
PT preventing atherosclerosis, hypertension, AIDS, allergy, multiple
PT sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
PT or hepatitis.
XX
PS Claim 1; SEQ ID NO 3; 347pp; English.
XX
CC The invention relates to a novel isolated polypeptide which is a human
CC kinase and phosphatase (KPP). The polypeptide of the invention
CC demonstrates cardiovascular, antiarteriosclerotic, hypotensive,
CC vasotropic, antiinflammatory, antianginal, anti-HIV, antiallergic,
CC antiasthmatic, immunosuppressive, antithyroid, dermatological,
CC antidiabetic, nephrotropic, antigout, gastrointestinal, neuroprotective,
CC osteopathic, antiarthritic, uropathic, ophthalmological, antirheumatic,
CC antiparkinsonian, nootropic, anticonvulsant, hepatotropic, antipsoriatic,
CC haemostatic, cytostatic, antilipaemic, antiparasitic, antihelmintic,
CC antibacterial, virucide, protozoacide and fungicide activities. The
CC kinase and phosphatase (KPP) polynucleotides, polypeptides, agonists and
CC antagonists may be useful for diagnosing, treating or preventing
CC disorders such as cardiovascular diseases, immune system disorders,
CC neurological disorders, disorders affecting growth and development, cell
CC proliferative disorders and viral, bacterial, fungal, parasitic,
CC protozoan or helminthic infections. Furthermore, the molecules of the
CC invention may be useful for creating transgenic animals to model human
CC disease and during gene therapy. The current sequence is that of a human
CC KPP protein of the invention.
XX
SQ Sequence 311 AA;

Query Match 77.3%; Score 34; DB 8; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTRRGL 8
|||||||
Db 175 RLTRRGL 181

RESULT 8
ADP29449
ID ADP29449 standard; protein; 933 AA.
XX
AC ADP29449;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #216.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 1447; 428pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
SQ Sequence 933 AA;

Query Match 77.3%; Score 34; DB 8; Length 933;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTRRGL 8
|||||||
Db 175 RLTRRGL 181

RESULT 9
ADE38441
ID ADE38441 standard; protein; 984 AA.
XX
AC ADE38441;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein 1420 amino acid sequence.
XX
KW tumourigenic disorder; angiogenic disorder; aberrant gene expression;
KW aberrant protein activity; cytostatic; antithyroid; antidiabetic;
KW ophthalmological; cancer; breast cancer; colon cancer; lung cancer;
KW prostatic cancer; Grave's disease; diabetic retinopathy; protein 1420.
XX
OS Homo sapiens.
XX
PN WO2003065006-A2.
XX
PD 07-AUG-2003.
XX
PF 30-JAN-2003; 2003WO-US002588.
XX
PR 31-JAN-2002; 2002US-0353600P.
PR 15-MAR-2002; 2002US-0364517P.
PR 09-APR-2002; 2002US-0371075P.
PR 10-APR-2002; 2002US-0371507P.
PR 16-APR-2002; 2002US-0372984P.
PR 19-APR-2002; 2002US-0374194P.
PR 24-MAY-2002; 2002US-0382995P.
PR 31-MAY-2002; 2002US-0385023P.
PR 14-JUN-2002; 2002US-0388853P.
PR 17-JUN-2002; 2002US-0389395P.
PR 25-JUN-2002; 2002US-0391324P.
PR 15-JUL-2002; 2002US-0395944P.
PR 22-JUL-2002; 2002US-0397726P.
PR 13-AUG-2002; 2002US-0403046P.
PR 22-AUG-2002; 2002US-0405155P.
PR 27-AUG-2002; 2002US-0406361P.
PR 25-OCT-2002; 2002US-0421195P.
PR 12-NOV-2002; 2002US-0425456P.
PR 19-NOV-2002; 2002US-0427626P.
PR 10-DEC-2002; 2002US-0432122P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Hunter JJ, Macbeth KJ, Tsai F, Lesoon A, Lightcap ES;
PI Williamson MW, Rudolph-Owen LA;
XX
DR WPI; 2003-646176/61.
DR N-PSDB; ADE38440.
XX
PT Treating subject having tumorigenic disorder or angiogenic disorder
PT caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic
PT acid, by administering a modulator.
XX
PS Disclosure; SEQ ID NO 102; 454pp; English.
XX
CC This invention relates to a novel method of treating a human subject
CC having a tumourigenic disorder or angiogenic disorder, caused by aberrant
CC gene expression or activity of an isolated protein, by administering a
CC modulator. The modulator may have cytostatic, antithyroid, antidiabetic
CC or ophthalmological activity. The method is useful for treating a subject
CC having a tumourigenic or angiogenic disorder, in particular for treating
CC cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC present sequence is the amino acid sequence of the novel isolated human
CC protein 1420 of the invention.
XX
SQ Sequence 984 AA;

Query Match 77.3%; Score 34; DB 7; Length 984;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTRRGL 8
|||||||
Db 175 RLTRRGL 181

RESULT 10
ADJ75552
ID ADJ75552 standard; protein; 984 AA.
XX
AC ADJ75552;
XX
DT 20-MAY-2004 (first entry)
XX
DE Marker gene related amino acid sequence SEQ ID NO:804.
XX
KW bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.

XX
OS Homo sapiens.
XX
PN EP1394274-A2.
XX
PD 03-MAR-2004.
XX
PF 04-AUG-2003; 2003EP-00254857.
XX
PR 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
PA (GENO-) GENOX RES INC.
XX
PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX
DR WPI; 2004-193155/19.
XX
PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX
PS Example 11; SEQ ID NO 804; 241pp; English.
XX
CC The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 984 AA;

Query Match 77.3%; Score 34; DB 8; Length 984;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTRRGL 8
||||||| |
Db 175 RLTRRGL 181

RESULT 11
AAY30684
ID AAY30684 standard; peptide; 10 AA.
XX
AC AAY30684;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 76.1%; Score 33.5; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLTR-RGLK 9
||||| ||||
Db 1 TRLTRARGLK 10

RESULT 12
AAY30683
ID AAY30683 standard; peptide; 10 AA.
XX
AC AAY30683;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.

OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 76.1%; Score 33.5; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLTR-RGLK 9
||||| ||||
Db 1 TRLTRDRGLK 10

RESULT 13
AAY30686
ID AAY30686 standard; peptide; 10 AA.
XX
AC AAY30686;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 76.1%; Score 33.5; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLTR-RGLK 9
||||| ||||
Db 1 TRLTRSRGLK 10

RESULT 14
AAY30682
ID AAY30682 standard; peptide; 10 AA.
XX
AC AAY30682;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX

DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 76.1%; Score 33.5; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLTR-RGLK 9
||||| ||||
Db 1 TRLTRERGLK 10

RESULT 15
AAY30685
ID AAY30685 standard; peptide; 10 AA.
XX
AC AAY30685;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

Query Match 76.1%; Score 33.5; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLTR-RGLK 9
||||| ||||
Db 1 TRLTRTRGLK 10

Search completed: December 29, 2004, 12:28:50
Job time : 54.9205 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 8.69318 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-13
Perfect score: 44
Sequence: 1 TRLTRRGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	77.3	984	1	A34076	protein-tyrosine k
2	34	77.3	1099	2	AE1065	conserved hypothet
3	33.5	76.1	596	2	S32802	apolipoprotein B -
4	33.5	76.1	4563	1	LPHUB	apolipoprotein B-1
5	33	75.0	189	2	B95329	probable ISRm25b t
6	33	75.0	274	2	D72044	hypothetical prote
7	33	75.0	274	2	A86581	hypothetical prote
8	33	75.0	368	2	AI1291	glycerol dehydroge
9	33	75.0	368	2	AG1663	glycerol dehydroge
10	33	75.0	427	2	A84155	hypothetical prote
11	33	75.0	451	2	F95869	probable ABC trans
12	33	75.0	514	2	F87592	hypothetical prote
13	33	75.0	633	2	T05005	hypothetical prote
14	33	75.0	871	2	T07863	probable polyprote
15	32	72.7	161	2	T35260	hypothetical prote
16	32	72.7	351	2	JC1175	hypothetical 38.5K
17	32	72.7	437	2	S54978	6-phosphofructokin
18	32	72.7	631	2	T29926	hypothetical prote
19	31	70.5	85	2	G81430	hypothetical prote
20	31	70.5	99	2	A87912	protein B0205.5 [i
21	31	70.5	128	2	F87342	conserved hypothet
22	31	70.5	155	2	E82452	anaerobic ribonucl
23	31	70.5	188	2	JU0451	hypothetical 21K p
24	31	70.5	188	2	I59116	myc protein - huma
25	31	70.5	188	2	A29867	hypothetical 20K p
26	31	70.5	188	2	I79500	myc protein - huma
27	31	70.5	248	2	S77172	glucose dehydrogen
28	31	70.5	261	2	G87325	siroheme synthase
29	31	70.5	275	2	S76916	hypothetical prote
30	31	70.5	279	2	A65020	hypothetical prote
31	31	70.5	318	2	B90399	probable acyl-coen
32	31	70.5	345	1	JH0185	D-amino-acid oxida
33	31	70.5	347	1	OXPGDA	D-amino-acid oxida
34	31	70.5	347	1	S01340	D-amino-acid oxida
35	31	70.5	347	1	JX0132	D-amino-acid oxida
36	31	70.5	357	2	S43067	hypothetical prote
37	31	70.5	364	2	C87455	alanine racemase [
38	31	70.5	393	2	H75444	branched-chain ami
39	31	70.5	393	2	AD2219	hypothetical prote
40	31	70.5	435	2	T03545	probable cobyrinic
41	31	70.5	462	2	T17948	ABC transporter pr
42	31	70.5	466	2	T41375	probable phosphogl
43	31	70.5	466	2	S62332	beta-fructofuranos
44	31	70.5	531	2	C75418	ribonucleoprotein
45	31	70.5	774	2	AI3372	malate dehydrogena

ALIGNMENTS

RESULT 1
A34076
protein-tyrosine kinase (EC 2.7.1.112) receptor type eph 1 precursor - human
N;Alternate names: receptor tyrosine kinase eph
C;Species: Homo sapiens (man)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: A34076; S44280
R;Hirai, H.; Maru, Y.; Hagiwara, K.; Nishida, J.; Takaku, F.
Science 238, 1717-1720, 1987
A;Title: A novel putative tyrosine kinase receptor encoded by the eph gene.
A;Reference number: A34076; MUID:88070650; PMID:2825356
A;Accession: A34076
A;Molecule type: mRNA
A;Residues: 1-984 <HIR>
A;Cross-references: UNIPROT:P21709; GB:M18391; NID:g339716; PIDN:AAA36747.1; PID:g33971
A;Note: the sequence in GenBank entry HUMTKR, release 111.0, has the codons GCG for 398
R;Tuzi, N.L.
submitted to the EMBL Data Library, November 1993
A;Description: An EGFR/eph chimeric receptor possesses ligand stimulated tyrosine kinas
A;Reference number: S44280
A;Accession: S44280
A;Molecule type: mRNA
A;Residues: 286-397,'A',399-580,'QRDRATDVDREDKLWLKPYVDLQAYEDPAQGALDF',583,625-984 <TUZ>
A;Cross-references: EMBL:Z27409; NID:g482916; PIDN:CAA81796.1; PID:g482917
C;Genetics:
A;Gene: GDB:EPHT1; EPH; EPHT
A;Cross-references: GDB:119875; OMIM:179610
A;Map position: 7q32-7q36
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat
C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-984/Product: protein-tyrosine kinase receptor type eph 1 #status predicted <MAT>
F;548-568/Domain: transmembrane #status predicted <TMM>
F;630-895/Domain: protein kinase homology <KIN>
F;638-646/Region: protein kinase ATP-binding motif
F;918-984/Domain: SAM homology <SAM>
F;59,338,414,478/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 77.3%; Score 34; DB 1; Length 984;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLTRRGL 8
              |||||||
Db        175 RLTRRGL 181
```

RESULT 2
AE1065
conserved hypothetical protein STY4851 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AE1065
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE1065
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1099 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD06970.1; PID:g16505611; GSPDB:GN00176
C;Genetics:
A;Gene: STY4851

Query Match 77.3%; Score 34; DB 2; Length 1099;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LTRRGLK 9
|||||||
Db 20 LTRRGLK 26

RESULT 3
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchic
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional re
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g93012
C;Superfamily: apolipoprotein B

Query Match 76.1%; Score 33.5; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 71;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLTR-RGLK 9
||||| ||||
Db 226 TRLTRKRGLK 235

RESULT 4
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A26
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Scc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,'
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNIP
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: re
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found
R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McCa
Nucleic Acids Res. 14, 7501-7503, 1986
A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q'
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',2
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428-
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQE
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; Ki
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipopr
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 reg
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180,
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yamana
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.

A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA
A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732
R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.;
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than on
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place c
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-
A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free s
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.,
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

Query Match 76.1%; Score 33.5; DB 1; Length 4563;
Best Local Similarity 90.0%; Pred. No. 4.7e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 TRLTR-RGLK 9
              ||||| ||||
Db       3385 TRLTRKRGLK 3394
```

RESULT 5
B95329
probable ISRm25b transposase [imported] - Sinorhizobium meliloti (strain 1021) magaplas
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: B95329
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: B95329
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-189 <KUR>
A;Cross-references: UNIPROT:Q92ZE8; GB:AE006469; PIDN:AAK65196.1; PID:g14523642; GSPDB:
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure

hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa0998
A;Genome: plasmid

Query Match 75.0%; Score 33; DB 2; Length 189;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTRRGLK 9
:|||||:
Db 106 KLTRRGLR 113

RESULT 6
D72044
hypothetical protein CP0022 [imported] - Chlamydophila pneumoniae (strains CWL029 and AR
N;Alternate names: hypothetical protein CPn0724
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: D72044; C81621
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: D72044
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-274 <ARN>
A;Cross-references: UNIPROT:Q9Z7I1; GB:AE001653; GB:AE001363; NID:g4377017; PIDN:AAD1886
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-274 <REA>
A;Cross-references: GB:AE002166; GB:AE002161; NID:g7188959; PIDN:AAF37918.1; PID:g718896
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CPn0724; CP0022
C;Superfamily: Chlamydia pneumoniae hypothetical protein CPn0724

Query Match 75.0%; Score 33; DB 2; Length 274;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRRGLK 9
|: |||||:
Db 41 TKATRRGLR 49

RESULT 7
A86581
hypothetical protein CPj0724 [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86581
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: A86581
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-274 <STO>
A;Cross-references: UNIPROT:Q9Z7I1; GB:BA000008; NID:g8979096; PIDN:BAA98931.1; GSPDB:GN
A;Experimental source: strain J138
C;Genetics:
A;Gene: CPj0724
C;Superfamily: Chlamydia pneumoniae hypothetical protein CPn0724

Query Match 75.0%; Score 33; DB 2; Length 274;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTRRGLK 9
|: |||||:
Db 41 TKATRRGLR 49

RESULT 8
AI1291
glycerol dehydrogenase homolog lmo1737 [imported] - Listeria monocytogenes (strain EGD-e
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C;Accession: AI1291
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1291
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <GLA>
A;Cross-references: UNIPROT:Q8Y6F0; GB:NC_003210; PIDN:CAC99815.1; PID:g16411191; GSPDB.
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1737
C;Superfamily: Glycerol dehydrogenase; lactaldehyde reductase homology

Query Match 75.0%; Score 33; DB 2; Length 368;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TRLTRRGLK 9
| | |||||
Db 27 THLERRGLK 35

RESULT 9
AG1663
glycerol dehydrogenase homolog lin1848 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C;Accession: AG1663
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1663
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <GLA>
A;Cross-references: UNIPROT:Q92AS1; GB:AL592022; PIDN:CAC97079.1; PID:g16414350; GSPDB:C
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1848
C;Superfamily: Glycerol dehydrogenase; lactaldehyde reductase homology

Query Match 75.0%; Score 33; DB 2; Length 368;
Best Local Similarity 77.8%; Pred. No. 58;

```
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              | | |||||
Db         27 THLERRGLK 35


RESULT 10
A84155
hypothetical protein BH4041 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A84155
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A84155
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <STO>
A;Cross-references: UNIPROT:Q9K5P7; GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB077
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH4041

  Query Match             75.0%;  Score 33;  DB 2;  Length 427;
  Best Local Similarity   66.7%;  Pred. No. 66;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||:|:|| |
Db        339 TRITKRGRK 347


RESULT 11
F95869
probable ABC transporter sugar-binding protein SMb20231 [imported] - Sinorhizobium melil
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95869
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: F95869
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-451 <KUR>
A;Cross-references: UNIPROT:Q92WV7; GB:AL591985; PIDN:CAC48622.1; PID:g15140094; GSPDB:G
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20231
A;Genome: plasmid

  Query Match             75.0%;  Score 33;  DB 2;  Length 451;
  Best Local Similarity   87.5%;  Pred. No. 70;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGL 8
              | ||||||
Db          9 TTLTRRGL 16


RESULT 12
F87592
hypothetical protein CC2774 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: F87592
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: F87592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-514 <STO>
A;Cross-references: UNIPROT:Q9A4Q6; GB:AE005673; NID:g13424372; PIDN:AAK24738.1; GSPDB:
C;Genetics:
A;Gene: CC2774

  Query Match             75.0%;  Score 33;  DB 2;  Length 514;
  Best Local Similarity   77.8%;  Pred. No. 78;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              | | |||||
Db        332 TELKRRGLK 340


RESULT 13
T05005
hypothetical protein T19P19.70 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05005
R;Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15394
A;Accession: T05005
A;Molecule type: DNA
A;Residues: 1-633 <BEV>
A;Cross-references: UNIPROT:O65655; EMBL:AL022605
A;Experimental source: cultivar Columbia; BAC clone T19P19
C;Genetics:
A;Map position: 4
A;Introns: 385/1; 448/1; 498/3
A;Note: T19P19.70

  Query Match             75.0%;  Score 33;  DB 2;  Length 633;
  Best Local Similarity   87.5%;  Pred. No. 95;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTRRGLK 9
              ||| ||||
Db         30 RLTTRGLK 37


RESULT 14
T07863
probable polyprotein - pineapple retrotransposon dea1 (fragment)
C;Species: Ananas comosus (pineapple)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T07863
R;Thomson, K.G.; Thomas, J.E.; Dietzgen, R.G.
Plant Mol. Biol. 38, 461-465, 1998
A;Title: Retrotransposon-like sequences integrated into the genome of pineapple, Ananas
A;Reference number: Z16184; MUID:98418625; PMID:9747853
A;Accession: T07863
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-871 <THO>
A;Cross-references: UNIPROT:O64892; EMBL:Y12432; NID:g2995404; PIDN:CAA73042.1; PID:g29
C;Genetics:
```

A;Mobile element: retrotransposon dea1

Query Match 75.0%; Score 33; DB 2; Length 871;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTRRGLK 9
              |||| :|:|
Db        272 TRLTHKGVK 280
```

RESULT 15
T35260
hypothetical protein SC5F2A.18 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35260
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21573
A;Accession: T35260
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-161 <OLI>
A;Cross-references: UNIPROT:Q9X7P1; EMBL:AL049587; PIDN:CAB40685.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC5F2A.18

Query Match 72.7%; Score 32; DB 2; Length 161;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLTRRGL 8
              ||||||:
Db        144 RLTRRGI 150
```

Search completed: December 29, 2004, 12:39:08
Job time : 9.69318 secs

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 52.5682 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-13
Perfect score: 44
Sequence: 1 TRLTRRGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	84.1	423	2	Q7WSQ9	Q7wsq9 arthrobacte
2	36	81.8	503	2	Q7P2I3	Q7p2i3 fusobacteri
3	36	81.8	506	2	Q8REI1	Q8rei1 fusobacteri
4	35	79.5	188	2	Q8QHI6	Q8qhi6 gallus gall
5	35	79.5	400	1	FXLE_MOUSE	Q8bid8 mus musculu
6	35	79.5	418	1	FXLE_HUMAN	Q8n1e6 homo sapien
7	35	79.5	581	2	Q740B8	Q740b8 mycobacteri
8	35	79.5	581	2	AAS03750	Aas03750 mycobacte
9	34	77.3	192	1	RAC4_HUMAN	O95916 homo sapien
10	34	77.3	321	2	Q98HM9	Q98hm9 rhizobium l
11	34	77.3	369	2	Q7RZ12	Q7rz12 neurospora
12	34	77.3	378	2	Q871B9	Q871b9 neurospora
13	34	77.3	439	1	COBB_RHILO	Q98kp1 rhizobium l
14	34	77.3	976	1	EPA1_HUMAN	P21709 homo sapien
15	34	77.3	1099	2	Q8Z0Z1	Q8z0z1 salmonella
16	34	77.3	1906	2	Q6C359	Q6c359 yarrowia li
17	33.5	76.1	414	2	Q7YQR5	Q7yqr5 aotus vocif
18	33.5	76.1	596	2	Q28473	Q28473 macaca fasc
19	33.5	76.1	3262	2	Q13788	Q13788 homo sapien
20	33.5	76.1	4563	1	APB_HUMAN	P04114 homo sapien
21	33.5	76.1	4563	2	Q7Z600	Q7z600 homo sapien
22	33	75.0	99	2	Q98ND4	Q98nd4 rhizobium l
23	33	75.0	184	2	Q73UW5	Q73uw5 mycobacteri
24	33	75.0	184	2	AAS05799	Aas05799 mycobacte
25	33	75.0	189	2	Q92ZE8	Q92ze8 rhizobium m
26	33	75.0	274	2	Q9Z7I1	Q9z7i1 chlamydia p
27	33	75.0	306	2	Q87WZ9	Q87wz9 pseudomonas
28	33	75.0	335	2	Q8PHW2	Q8phw2 xanthomonas
29	33	75.0	367	2	Q71YT1	Q71yt1 listeria mo
30	33	75.0	367	2	AAT04533	Aat04533 listeria
31	33	75.0	368	2	Q92AS1	Q92as1 listeria in
32	33	75.0	368	2	Q8Y6F0	Q8y6f0 listeria mo
33	33	75.0	383	2	Q9RDB7	Q9rdb7 streptomyce
34	33	75.0	398	2	Q82CW6	Q82cw6 streptomyce
35	33	75.0	423	2	Q75TW9	Q75tw9 bacillus ha
36	33	75.0	423	2	BAD18188	Bad18188 bacillus
37	33	75.0	427	2	Q75TN6	Q75tn6 bacillus fi
38	33	75.0	427	2	Q75TQ1	Q75tq1 bacillus al
39	33	75.0	427	2	Q75TX1	Q75tx1 bacillus ha
40	33	75.0	427	2	Q75TX3	Q75tx3 bacillus ha
41	33	75.0	427	2	Q75TX6	Q75tx6 bacillus ha
42	33	75.0	427	2	Q75TX7	Q75tx7 bacillus ha
43	33	75.0	427	2	Q75TZ5	Q75tz5 bacillus ha
44	33	75.0	427	2	Q75TZ7	Q75tz7 bacillus ha
45	33	75.0	427	2	Q75TZ8	Q75tz8 bacillus ha

ALIGNMENTS

RESULT 1
Q7WSQ9
ID Q7WSQ9 PRELIMINARY; PRT; 423 AA.
AC Q7WSQ9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transporter protein.
OS Arthrobacter ilicis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=43665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rue61a;
RX MEDLINE=22753791; PubMed=12730200;
RA Parschat K., Hauer B., Kappl R., Kraft R., Huettermann J., Fetzner S.;
RT "Gene Cluster of Arthrobacter ilicis R.61a Involved in the Degradation
RT of Quinaldine to Anthranilate. Characterization and Functional
RT Expression of the Quinaldine 4-oxidase qoxLMS Genes.";
RL J. Biol. Chem. 278:27483-27494(2003).
DR EMBL; AJ537472; CAD61041.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
SQ SEQUENCE 423 AA; 43696 MW; BB11CBADA85DF241 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 423;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTRRGLK 9
              ||||::|||
Db        207 TRLTKQGLK 215
```

RESULT 2
Q7P2I3
ID Q7P2I3 PRELIMINARY; PRT; 503 AA.
AC Q7P2I3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical cytosolic protein.
GN Name=FNV0795;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacterales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49256;

```
RA   Karpatral V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA   Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA   Haselkorn R., Overbeek R., Kyrpides N.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABF01000083; EAA23797.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   503 AA;  57362 MW;  70DE5CADE118516C CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 503;
  Best Local Similarity   77.8%;  Pred. No. 49;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||| |||:|
Db        126 TRLERRGIK 134


RESULT 3
Q8REI1
ID   Q8REI1      PRELIMINARY;      PRT;   506 AA.
AC   Q8REI1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical cytosolic protein.
GN   OrderedLocusNames=FN1121;
OS   Fusobacterium nucleatum (subsp. nucleatum).
OC   Bacteria; Fusobacteria; Fusobacterales; Fusobacteriaceae;
OC   Fusobacterium.
OX   NCBI_TaxID=76856;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 25586;
RX   MEDLINE=21886394; PubMed=11889109;
RA   Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA   Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA   Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA   Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA   Fonstein M., Kyrpides N.C., Overbeek R.;
RT   "Genome sequence and analysis of the oral bacterium Fusobacterium
RT   nucleatum strain ATCC 25586.";
RL   J. Bacteriol. 184:2005-2018(2002).
DR   EMBL; AE010616; AAL95317.1; -.
DR   GO; GO:0008233; F:peptidase activity; IEA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   506 AA;  57758 MW;  CE8577579504D47F CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 506;
  Best Local Similarity   77.8%;  Pred. No. 49;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||| |||:|
Db        126 TRLERRGIK 134


RESULT 4
Q8QHI6
ID   Q8QHI6      PRELIMINARY;      PRT;   188 AA.
AC   Q8QHI6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   PPA (Fragment).
GN   Name=PPA;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21972450; PubMed=11976951;
RA   Das T., Purkayastha-Mukherjee C., D'Angelo J., Weir M.;
RT   "A conserved F-box gene with unusual transcript localization.";
RL   Dev. Genes Evol. 212:134-140(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Das T.K., Purkayastha-Mukherjee C., D'Angelo J., Weir M.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF467464; AAL75968.1; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007089; LRR_cys.
DR   Pfam; PF00560; LRR; 4.
FT   NON_TER       1      1
FT   NON_TER     188    188
SQ   SEQUENCE   188 AA;  20629 MW;  21702832DA5CE865 CRC64;

  Query Match             79.5%;  Score 35;  DB 2;  Length 188;
  Best Local Similarity   66.7%;  Pred. No. 27;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||:|:|||:
Db        161 TRITKRGLE 169


RESULT 5
FXLE_MOUSE
ID   FXLE_MOUSE     STANDARD;      PRT;   400 AA.
AC   Q8BID8; Q8R5H7; Q8VDT7; Q922N5;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   F-box/LRR-repeat protein 14 (F-box and leucine-rich repeat protein
DE   14).
GN   Name=Fbxl14; Synonyms=Ppa;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Breast tumor, and Heart;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
```

```
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 207-390 FROM N.A.
RX   MEDLINE=21972450; PubMed=11976951; DOI=10.1007/s00427-002-0222-7;
RA   Das T., Purkayastha-Mukherjee C., D'Angelo J., Weir M.;
RT   "A conserved F-box gene with unusual transcript localization.";
RL   Dev. Genes Evol. 212:134-140(2002).
CC   -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC       proteins and promotes their ubiquitination and degradation.
CC   -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 F-box domain.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AK084506; BAC39201.1; -.
DR   EMBL; BC006913; AAH06913.1; -.
DR   EMBL; BC021329; AAH21329.1; -.
DR   EMBL; AF467463; AAL75967.1; -.
DR   MGD; MGI:2141676; Fbxl14.
DR   InterPro; IPR001810; F-box.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007089; LRR_cys.
DR   InterPro; IPR008945; Skp1_Skp2.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00560; LRR; 6.
DR   SMART; SM00256; FBOX; 1.
DR   PROSITE; PS50181; FBOX; FALSE_NEG.
KW   Leucine-rich repeat; Repeat; Ubl conjugation pathway.
FT   DOMAIN       2     48       F-box.
FT   REPEAT      91    120       LRR 1.
FT   REPEAT     170    194       LRR 2.
FT   REPEAT     203    231       LRR 3.
FT   REPEAT     254    280       LRR 4.
FT   REPEAT     331    356       LRR 5.
FT   REPEAT     357    381       LRR 6.
FT   CONFLICT    22     22       V -> F (in Ref. 2; AAH21329).
SQ   SEQUENCE   400 AA;  43864 MW;  E0B297E4B4F83C22 CRC64;

  Query Match           79.5%;  Score 35;  DB 1;  Length 400;
  Best Local Similarity   66.7%;  Pred. No. 63;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||:|:|||:
Db        367 TRITKRGLE 375


RESULT 6
FXLE_HUMAN
ID   FXLE_HUMAN     STANDARD;      PRT;   418 AA.
AC   Q8N1E6;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   F-box/LRR-repeat protein 14 (F-box and leucine-rich repeat protein
DE   14).
GN   Name=FBXL14;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC       proteins and promotes their ubiquitination and degradation.
CC   -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 F-box domain.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; BC028132; AAH28132.1; -.
DR   Genew; HGNC:28624; FBXL14.
DR   InterPro; IPR001810; F-box.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007089; LRR_cys.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00560; LRR; 6.
DR   SMART; SM00256; FBOX; 1.
DR   PROSITE; PS50181; FBOX; FALSE_NEG.
KW   Leucine-rich repeat; Repeat; Ubl conjugation pathway.
FT   DOMAIN       2     48       F-box.
FT   REPEAT      91    120       LRR 1.
FT   REPEAT     170    194       LRR 2.
FT   REPEAT     203    231       LRR 3.
FT   REPEAT     254    280       LRR 4.
```

```
FT   REPEAT       331    356       LRR 5.
FT   REPEAT       357    381       LRR 6.
SQ   SEQUENCE   418 AA;  45886 MW;  5779961C8177779F CRC64;

  Query Match             79.5%;  Score 35;  DB 1;  Length 418;
  Best Local Similarity   66.7%;  Pred. No. 66;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRGLK 9
              ||:|:|||:
Db        367 TRITKRGLE 375


RESULT 7
Q740B8
ID   Q740B8      PRELIMINARY;      PRT;   581 AA.
AC   Q740B8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=MAP1433c;
OS   Mycobacterium paratuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium avium complex (MAC).
OX   NCBI_TaxID=1770;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=k10;
RA   Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017232; AAS03750.1; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   581 AA;  61736 MW;  4082D4B95CB496B0 CRC64;

  Query Match             79.5%;  Score 35;  DB 2;  Length 581;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRG 7
              |||||||
Db        243 TRLTRRG 249


RESULT 8
AAS03750
ID   AAS03750    PRELIMINARY;      PRT;   581 AA.
AC   AAS03750;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   MAP1433C.
OS   Mycobacterium paratuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1770;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=k10;
RA   Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017232; AAS03750.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   581 AA;  61736 MW;  4082D4B95CB496B0 CRC64;

  Query Match             79.5%;  Score 35;  DB 2;  Length 581;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTRRG 7
              |||||||
Db        243 TRLTRRG 249


RESULT 9
RAC4_HUMAN
ID   RAC4_HUMAN     STANDARD;      PRT;   192 AA.
AC   O95916;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Ras-related C3 botulinum toxin substrate 4 (p21-Rac4).
GN   Name=RAC4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pearce A.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the small GTPase superfamily. Rho family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL022576; -; NOT_ANNOTATED_CDS.
DR   HSSP; P15154; 1HH4.
DR   Genew; HGNC:31113; RAC4.
DR   InterPro; IPR003578; GTPase_Rho.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; Ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00174; RHO; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Lipoprotein; Polymorphism; Prenylation.
FT   NP_BIND      10     17       GTP (By similarity).
FT   NP_BIND      57     61       GTP (By similarity).
FT   NP_BIND     115    118       GTP (By similarity).
FT   DOMAIN       32     40       Effector region (Potential).
FT   LIPID       189    189       S-geranylgeranyl cysteine.
FT   VARIANT      14     14       V -> A (in dbSNP:5833).
FT                                /FTId=VAR_014551.
FT   VARIANT      27     27       A -> T (in dbSNP:5824).
FT                                /FTId=VAR_014552.
FT   VARIANT      29     29       P -> S (in dbSNP:5827).
FT                                /FTId=VAR_014553.
SQ   SEQUENCE   192 AA;  21383 MW;  09C5DFE64C8E6053 CRC64;

  Query Match             77.3%;  Score 34;  DB 1;  Length 192;
  Best Local Similarity   100.0%;  Pred. No. 47;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTRRGLK 9
              |||||||
Db        160 LTRRGLK 166


RESULT 10
Q98HM9
ID   Q98HM9      PRELIMINARY;      PRT;   321 AA.
AC   Q98HM9;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mll2796 protein.
GN   OrderedLocusNames=mll2796;
OS   Rhizobium loti (Mesorhizobium loti).
```

```
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082936; PubMed=11214974;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti (supplement).";
RL   DNA Res. 7:381-406(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003000; BAB49837.1; -.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR004843; M-pesterase.
DR   Pfam; PF00149; Metallophos; 1.
KW   Complete proteome.
SQ   SEQUENCE   321 AA;  34893 MW;  C6FB1DA0CC1D73D8 CRC64;

  Query Match             77.3%;  Score 34;  DB 2;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 82;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTRRGL 8
              |||||||
Db        289 RLTRRGL 295


RESULT 11
Q7RZ12
ID   Q7RZ12      PRELIMINARY;      PRT;   369 AA.
AC   Q7RZ12;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Predicted protein.
GN   Name=NCU07194.1;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000721; EAA28183.1; -.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR000733; Flav_monooxygnse.
DR   InterPro; IPR003042; Rng_mnoxygenase.
DR   Pfam; PF01360; Monooxygenase; 1.
DR   PRINTS; PR00420; RNGMNOXGNASE.
SQ   SEQUENCE   369 AA;  40627 MW;  D53DE9368557BE47 CRC64;

  Query Match             77.3%;  Score 34;  DB 2;  Length 369;
  Best Local Similarity   75.0%;  Pred. No. 96;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTRRGLK 9
              |:|||||:
Db         63 RITRRGLR 70


RESULT 12
Q871B9
ID   Q871B9      PRELIMINARY;      PRT;   378 AA.
AC   Q871B9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Related to oxygenase.
GN   Name=B8G12.280;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX294027; CAD71088.1; -.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR000733; Flav_monooxygnse.
DR   InterPro; IPR003042; Rng_mnoxygenase.
DR   Pfam; PF01360; Monooxygenase; 1.
DR   PRINTS; PR00420; RNGMNOXGNASE.
SQ   SEQUENCE   378 AA;  41711 MW;  405EB2FB76BCA5A0 CRC64;

  Query Match             77.3%;  Score 34;  DB 2;  Length 378;
  Best Local Similarity   75.0%;  Pred. No. 98;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTRRGLK 9
              |:|||||:
Db         63 RITRRGLR 70


RESULT 13
COBB_RHILO
ID   COBB_RHILO      STANDARD;      PRT;   439 AA.
AC   Q98KP1;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Cobyrinic acid A,C-diamide synthase.
```

```
GN   Name=cobB; OrderedLocusNames=mlr1387;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
CC   -!- FUNCTION: Responsible for the amidation of carboxylic groups at
CC       position A and C of either cobyrinic acid or hydrogenobrynic acid.
CC       NH(2) groups are provided by glutamine, and one molecule of ATP is
CC       hydrogenolyzed for each amidation (By similarity).
CC   -!- PATHWAY: Cobalamin biosynthesis.
CC   -!- SIMILARITY: Belongs to the cobB/cobQ family. CobB subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP002997; BAB48773.1; -.
DR   HAMAP; MF_00027; -; 1.
DR   InterPro; IPR002586; CbiA_P.
DR   InterPro; IPR004484; CbiA_synth.
DR   Pfam; PF01656; CbiA; 1.
DR   TIGRFAMs; TIGR00379; cobB; 1.
KW   Cobalamin biosynthesis; Complete proteome.
SQ   SEQUENCE   439 AA;  45879 MW;  DCFB9F997D2948B8 CRC64;

  Query Match             77.3%;  Score 34;  DB 1;  Length 439;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTRRGLK 9
              |||||||
Db         28 LTRRGLK 34


RESULT 14
EPA1_HUMAN
ID   EPA1_HUMAN     STANDARD;      PRT;   976 AA.
AC   P21709; Q15405;
DT   01-MAY-1991 (Rel. 18, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Ephrin type-A receptor 1 precursor (EC 2.7.1.112) (Tyrosine-protein
DE   kinase receptor EPH).
GN   Name=EPHA1; Synonyms=EPHT1, EPHT, EPH;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88070650; PubMed=2825356;
RA   Hirai H., Maru Y., Hagiwara K., Nishida J., Takaku F.;
RT   "A novel putative tyrosine kinase receptor encoded by the eph gene.";
RL   Science 238:1717-1720(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99299440; PubMed=10369740;
RA   Owshalimpur D., Kelley M.J.;
RT   "Genomic structure of the EPHA1 receptor tyrosine kinase gene.";
RL   Mol. Cell. Probes 13:169-173(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA   Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA   Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA   Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA   Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA   Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA   Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA   Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA   Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA   Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA   Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA   Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA   Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
RA   Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA   Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA   Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA   Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubb K.,
RA   Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA   Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA   Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA   Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA   Waterston R.H., Wilson R.K.;
RT   "The DNA sequence of human chromosome 7.";
RL   Nature 424:157-164(2003).
RN   [4]
RP   SEQUENCE OF 286-976 FROM N.A.
RC   TISSUE=Placenta;
RA   Tuzi N.L.;
RL   Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for members of the ephrin-A family. Binds with
CC       a low affinity to ephrin-A1.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC       tyrosine phosphate.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Overexpressed in several carcinomas.
CC   -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Ephrin
CC       receptor subfamily.
CC   -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M18391; AAA36747.1; ALT_SEQ.
DR   EMBL; AF101171; AAD43440.1; -.
DR   EMBL; AF101165; AAD43440.1; JOINED.
DR   EMBL; AF101166; AAD43440.1; JOINED.
DR   EMBL; AF101167; AAD43440.1; JOINED.
DR   EMBL; AF101168; AAD43440.1; JOINED.
DR   EMBL; AF101169; AAD43440.1; JOINED.
DR   EMBL; AF101170; AAD43440.1; JOINED.
DR   EMBL; AC092214; AAS07458.1; -.
DR   EMBL; Z27409; CAA81796.1; -.
DR   PIR; A34076; A34076.
DR   HSSP; P54763; 1JPA.
DR   Genew; HGNC:3385; EPHA1.
DR   MIM; 179610; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . .; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001090; Ephrin_receptor.
DR   InterPro; IPR003961; FN_III.
```

```
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR008979; Gal_bind_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR001660; SAM.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   InterPro; IPR008266; Tyr_pkinase_AS.
DR   InterPro; IPR001426; YKase_receptorV.
DR   Pfam; PF01404; Ephrin_lbd; 1.
DR   Pfam; PF00041; fn3; 2.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD001495; Ephrin_receptor; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00615; EPH_lbd; 1.
DR   SMART; SM00060; FN3; 2.
DR   SMART; SM00454; SAM; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PROSITE; PS50853; FN3; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   ATP-binding; Glycoprotein; Phosphorylation; Receptor; Repeat; Signal;
KW   Transferase; Transmembrane; Tyrosine-protein kinase.
FT   SIGNAL        1     23       Potential.
FT   CHAIN        24    976       Ephrin type-A receptor 1.
FT   DOMAIN       24    547       Extracellular (Potential).
FT   TRANSMEM    548    568       Potential.
FT   DOMAIN      569    976       Cytoplasmic (Potential).
FT   DOMAIN      191    329       Cys-rich.
FT   DOMAIN      332    437       Fibronectin type-III 1.
FT   DOMAIN      446    535       Fibronectin type-III 2.
FT   DOMAIN      624    884       Protein kinase.
FT   DOMAIN      913    976       SAM.
FT   SITE        974    976       PDZ-binding motif (Potential).
FT   NP_BIND     630    638       ATP (By similarity).
FT   BINDING     656    656       ATP (By similarity).
FT   ACT_SITE    749    749       By similarity.
FT   MOD_RES     599    599       Phosphotyrosine (by autocatalysis)
FT                                (Potential).
FT   MOD_RES     605    605       Phosphotyrosine (by autocatalysis)
FT                                (Potential).
FT   MOD_RES     781    781       Phosphotyrosine (by autocatalysis)
FT                                (Potential).
FT   MOD_RES     930    930       Phosphotyrosine (by autocatalysis)
FT                                (Potential).
FT   CARBOHYD    338    338       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    414    414       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    478    478       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    160    160       A -> V (in Ref. 3).
FT   CONFLICT    398    398       G -> A (in Ref. 2 and 4).
FT   CONFLICT    581    616       QRDRATDVDREDKLWLKPYVDLQAYEDPAQGALDFT -> H
FT                                VTAPPMWIERTSCAEALCGTSRHTRTLHREPWTLPGGWSNF
FT                                PS (in Ref. 1).
FT   CONFLICT    900    900       V -> M (in Ref. 3).
SQ   SEQUENCE   976 AA;  108066 MW;  9F8F2D6A8E683D95 CRC64;

  Query Match             77.3%;  Score 34;  DB 1;  Length 976;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTRRGL 8
              |||||||
Db        175 RLTRRGL 181


RESULT 15
Q8Z0Z1
ID   Q8Z0Z1      PRELIMINARY;      PRT;  1099 AA.
AC   Q8Z0Z1; Q7C4Y0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein STY4851.
GN   OrderedLocusNames=STY4851, t4545;
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
DR   EMBL; AL627283; CAD06970.1; -.
DR   EMBL; AE016849; AAO71983.1; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   1099 AA;  125223 MW;  6131DD4E8AB5F6A5 CRC64;

  Query Match             77.3%;  Score 34;  DB 2;  Length 1099;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LTRRGLK 9
              |||||||
Db         20 LTRRGLK 26


Search completed: December 29, 2004, 12:37:38
Job time : 54.6793 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:10:41 ; Search time 54.9205 Seconds
(without alignments)
58.786 Million cell updates/sec

Title: US-09-823-418-14
Perfect score: 44
Sequence: 1 TRLTKRGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	9	2	AAY30695	Aay30695 Apo-B100
2	41	93.2	9	2	AAY30694	Aay30694 Apo-B100
3	38	86.4	9	2	AAY30696	Aay30696 Apo-B100
4	38	86.4	404	6	ABR43240	Abr43240 Human PMM
5	38	86.4	548	5	ABG97506	Abg97506 Human NOV
6	35	79.5	146	3	AAG40845	Aag40845 Zea mays
7	34	77.3	238	8	ADN47506	Adn47506 Thermococ
8	34	77.3	335	7	ADG30649	Adg30649 Xanthomon
9	34	77.3	840	6	ABU48780	Abu48780 Protein e
10	34	77.3	3079	2	AAR59926	Aar59926 GAP prote
11	33.5	76.1	10	2	AAY30689	Aay30689 Apo-B100
12	33.5	76.1	10	2	AAY30688	Aay30688 Apo-B100
13	33.5	76.1	11	2	AAW57205	Aaw57205 Apo B bin
14	33.5	76.1	13	2	AAW57207	Aaw57207 Apo B 100
15	33.5	76.1	15	2	AAW41261	Aaw41261 Apolipopr
16	33.5	76.1	15	2	AAW96892	Aaw96892 ApoB-100
17	33.5	76.1	20	6	ABJ37575	Abj37575 Heparin b
18	33.5	76.1	22	2	AAW57208	Aaw57208 Apo B 100
19	33.5	76.1	22	2	AAW57209	Aaw57209 Apo B 100
20	33.5	76.1	34	5	AAE14541	Aae14541 Human apo
21	33.5	76.1	36	2	AAW96876	Aaw96876 Nucleic a
22	33.5	76.1	37	2	AAW64587	Aaw64587 Human apo
23	33.5	76.1	51	2	AAW96845	Aaw96845 Nucleic a
24	33.5	76.1	343	4	ABB37687	Abb37687 Peptide #
25	33.5	76.1	343	4	ABG52504	Abg52504 Human liv
26	33.5	76.1	377	2	AAR72704	Aar72704 Human apo
27	33.5	76.1	377	2	AAR34031	Aar34031 Sequence
28	33.5	76.1	2463	8	ADJ57400	Adj57400 Human apo
29	33.5	76.1	3923	2	AAY31237	Aay31237 Human Apo
30	33.5	76.1	4536	2	AAW41262	Aaw41262 Apolipopr
31	33.5	76.1	4536	2	AAW96826	Aaw96826 Amino aci
32	33.5	76.1	4560	5	AAU98981	Aau98981 Human apo
33	33.5	76.1	4561	7	ADD48677	Add48677 Human Pro
34	33.5	76.1	4563	5	AAO15893	Aao15893 Human apo
35	33.5	76.1	4563	6	ABR40253	Abr40253 Human ali
36	33.5	76.1	4563	6	ABU79140	Abu79140 Apolipopr
37	33.5	76.1	4563	7	ADF43408	Adf43408 Apolipopr
38	33.5	76.1	4563	8	ADH18871	Adh18871 Human apo
39	33.5	76.1	4563	8	ADH18870	Adh18870 Human apo
40	33.5	76.1	4563	8	ADO33445	Ado33445 Human apo
41	33.5	76.1	4563	8	ADO33447	Ado33447 Human apo
42	33.5	76.1	4590	4	AAU33184	Aau33184 Novel hum
43	33	75.0	11	2	AAY30699	Aay30699 Apo-B100
44	33	75.0	11	2	AAY30697	Aay30697 Apo-B100
45	33	75.0	120	4	AAE12587	Aae12587 Pseudomon

ALIGNMENTS

RESULT 1
AAY30695
ID AAY30695 standard; peptide; 9 AA.
XX
AC AAY30695;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can

```
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 44;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              |||||||||
Db          1 TRLTKRGLK 9


RESULT 2
AAY30694
ID   AAY30694 standard; peptide; 9 AA.
XX
AC   AAY30694;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 9 AA;

  Query Match             93.2%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.7e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||||:||||
Db          1 TRLTRRGLK 9


RESULT 3
AAY30696
ID   AAY30696 standard; peptide; 9 AA.
XX
AC   AAY30696;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 9 AA;

  Query Match             86.4%;  Score 38;  DB 2;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
```

```
              ||||::|||
Db          1 TRLTRKGLK 9
```

RESULT 4
ABR43240
ID ABR43240 standard; protein; 404 AA.
XX
AC ABR43240;
XX
DT 07-JUL-2003 (first entry)
XX
DE Human PMMM-1 protein SEQ ID NO:1.
XX
KW Human; protein modification and maintenance molecule; PMMM; cytostatic;
KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective; AIDS;
KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory; cancer;
KW thyromimetic; gene therapy; cell proliferative disorder; atherosclerosis;
KW neurological disorder; epilepsy; Huntington's disease; stroke; allergy;
KW immune disorder; inflammatory disorder; developmental disorder;
KW hypothyroidism; Cushing's syndrome; infection.
XX
OS Homo sapiens.
XX
PN WO2003025131-A2.
XX
PD 27-MAR-2003.
XX
PF 13-SEP-2002; 2002WO-US029221.
XX
PR 14-SEP-2001; 2001US-0322196P.
PR 21-SEP-2001; 2001US-0324134P.
PR 05-OCT-2001; 2001US-0327233P.
PR 26-OCT-2001; 2001US-0346198P.
PR 02-NOV-2001; 2001US-0343980P.
PR 09-NOV-2001; 2001US-0348887P.
PR 16-NOV-2001; 2001US-0332423P.
PR 28-NOV-2001; 2001US-0334145P.
PR 28-NOV-2001; 2001US-0334229P.
PR 06-DEC-2001; 2001US-0337451P.
PR 25-JAN-2002; 2002US-0351928P.
PR 21-MAR-2002; 2002US-0366837P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sprague WW, Chawla NK, Warren BA, Tang YT, Elliott VS;
PI Marquis JP, Li JX, Griffin JA, Gietzen KJ, Yang J, Lu DAM;
PI Emerling BM, Duggan BM, Richardson TW, Lee SY, Ramkumar J, Becha SD;
PI Lehr-Mason PM, Swarnakar A, Tran UK, Kable AE, Hafalia AJA, Khare R;
XX
DR WPI; 2003-354597/33.
DR N-PSDB; ACC59959.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, AIDS, epilepsy, or
PT infections.
XX
PS Claim 1; Page 206-207; 270pp; English.
XX
CC ACC59959 to ACC59989 encode the human protein modification and
CC maintenance molecule proteins given in ABR43240 to ABR43270, designated
CC PMMM-1 to PMMM-31 (I). (I) have cytostatic, antiarteriosclerotic,
CC anticonvulsant, nootropic, neuroprotective, cerebroprotective, anti-HIV,
CC antiallergic, antiinflammatory and thyromimetic activities, and can be
CC used in gene therapy. The PMMM polypeptides and polynucleotides are
CC useful in diagnosing, treating and preventing diseases or conditions
CC associated with the decreased expression or overexpression of PMMM, such
CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)
CC disorders, or infections. They are also useful in assessing the effects
CC of exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of PMMM. The PMMMs or their fragments are useful in screening
CC compounds for effectiveness as agonist or antagonist of the polypeptides,
CC or in altering the expression of the target polynucleotide and compounds
CC that specifically bind to or modulate the activity of the polypeptide.
CC The microarray is useful in monitoring or measuring protein-protein
CC interactions, drug-target interactions, and gene expression profiles
XX
SQ Sequence 404 AA;

```
  Query Match             86.4%;  Score 38;  DB 6;  Length 404;
  Best Local Similarity   77.8%;  Pred. No. 20;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||||:
Db        367 TRITKRGLE 375
```

RESULT 5
ABG97506
ID ABG97506 standard; protein; 548 AA.
XX
AC ABG97506;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human NOVX25 protein.
XX
KW Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
KW tuberous sclerosis; hypercalcaemia; Parkinson's disease; depression;
KW Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
KW scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
KW asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
KW bacterial infection; parasitic infection; graft-versus-host disease;
KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
KW angiogenesis.
XX
OS Homo sapiens.
XX
PN WO200272770-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US007283.
XX
PR 08-MAR-2001; 2001US-0274281P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-MAY-2001; 2001US-0288148P.
PR 31-MAY-2001; 2001US-0294821P.
PR 31-OCT-2001; 2001US-0335302P.
PR 04-DEC-2001; 2001US-0338375P.
PR 07-MAR-2002; 2002US-00094466.
XX
PA (CURA-) CURAGEN CORP.

XX
PI Spytek KA, Vernet CA, Tchernev VT, Malyankar UM, Gerlach VL;
PI Li L, Zerhusen BD, Patturajan M, Gusev VY, Kekuda R, Pena CEA;
PI Zhong M, Gangolli EA, Taupier RJ;
XX
DR WPI; 2002-713508/77.
DR N-PSDB; ABS78750.
XX
PT New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
PT Parkinson's disease.
XX
PS Claim 1; Page 161; 266pp; English.
XX
CC The present invention relates to a new polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease,
CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
CC polypeptides and antibodies are useful for treating, preventing or
CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
CC syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
CC scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
CC infections, or graft-versus-host disease. The nucleic acids and
CC polypeptides may also be used as targets for the identification of small
CC molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and
CC angiogenesis, in gene therapy, in generation of antibodies that bind
CC immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids are further used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine, and
CC pharmacogenomics. The present amino acid sequence represents a human NOVX
CC protein of the invention
XX
SQ Sequence 548 AA;

Query Match 86.4%; Score 38; DB 5; Length 548;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTKRGLK 9
              ||:|||||:
Db        515 TRITKRGLE 523
```

RESULT 6
AAG40845
ID AAG40845 standard; protein; 146 AA.
XX
AC AAG40845;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 50735.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX

PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 79.5%; Score 35; DB 3; Length 146;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 RLTKRGLK 9
              | ||||||
Db         44 RFTKRGLK 51
```

RESULT 7
ADN47506
ID ADN47506 standard; protein; 238 AA.
XX
AC ADN47506;
XX
DT 01-JUL-2004 (first entry)
XX
DE Thermococcus kodakaraensis KOD1 protein sequence SeqID1384.
XX
KW gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaebacterium; KOD1;
KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology.
XX
OS Thermococcus kodakaraensis.
XX
PN WO2004022736-A1.
XX
PD 18-MAR-2004.
XX
PF 29-AUG-2003; 2003WO-IB003597.
XX
PR 30-AUG-2002; 2002JP-00319011.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Imanaka T, Atomi H;
XX

DR WPI; 2004-257583/24.
XX
PT Method for disrupting targeted gene in genome of organism particularly
PT thermostable bacterium and with genome chips for analysis, applicable in
PT studying gene structure and functions.
XX
PS Claim 9; SEQ ID NO 1384; 598pp; Japanese.
XX
CC This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene,
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaebacterium, particularly
CC Thermococcus kodakaraensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakaraensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 238 AA;

Query Match 77.3%; Score 34; DB 8; Length 238;
Best Local Similarity 77.8%; Pred. No. 78;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTKRGLK 9
||||| |:|
Db 62 TRLTKIGIK 70

RESULT 8
ADG30649
ID ADG30649 standard; protein; 335 AA.
XX
AC ADG30649;
XX
DT 26-FEB-2004 (first entry)
XX
DE Xanthomonas axonopodis pv citri plant pathology-related XAC3136 protein.
XX
KW Xanthomonas microorganism; plant; pathology; bacterial pest; Xac; Xcc;
KW XAC.
XX
OS Xanthomonas axonopodis pv. citri.
XX
PN WO2003089647-A1.
XX
PD 30-OCT-2003.
XX
PF 22-APR-2003; 2003WO-BR000060.
XX
PR 22-APR-2002; 2002US-0374620P.
XX
PA (AMPA-) FUNDACAO AMPARO A PESQUISA DO ESTADO.
XX
PI Da Silva ACR, Farah SC, Quaggio RB, Reinach FDC, Ferro JA;
PI De Oliveira JCF, De Laia ML, Setubal JC, Furlan LR;
XX
DR WPI; 2003-865444/80.
DR N-PSDB; ADG30648.
XX
PT New nucleic acid molecule from a Xanthomonas microorganism, useful in
PT determining the presence of Xanthomonas bacteria in a sample.
XX
PS Claim 8; SEQ ID NO 6; 145pp; English.
XX
CC The invention relates to a novel isolated nucleic acid molecule from a
CC Xanthomonas microorganism where the nucleic acid molecule is associated
CC with pathogenicity caused by the Xanthomonas microorganism, or its
CC variant, that causes reduced or enhanced pathogenicity. The nucleic acid
CC of the invention may be useful in detecting the presence of Xanthomonas
CC bacteria in a sample, as well as in plant pathology, for example, for
CC identifying nucleic acid molecules and proteins involved in pathology
CC caused by bacterial pests. The current sequence is that of the
CC Xanthomonas axonopodis pv. citri (Xac) plant pathology-related XAC
CC protein of the invention.
XX
SQ Sequence 335 AA;

Query Match 77.3%; Score 34; DB 7; Length 335;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTKRGLK 9
||| :||||
Db 294 TRLLERGLK 302

RESULT 9
ABU48780
ID ABU48780 standard; protein; 840 AA.
XX
AC ABU48780;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #34307.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Ureaplasma urealyticum.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA52650.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 76704; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular

```
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 840 AA;

  Query Match             77.3%;  Score 34;  DB 6;  Length 840;
  Best Local Similarity   87.5%;  Pred. No. 2.9e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTKRGLK 9
              |||||| |
Db        744 RLTKRGAK 751


RESULT 10
AAR59926
ID   AAR59926 standard; protein; 3079 AA.
XX
AC   AAR59926;
XX
DT   25-MAR-2003  (revised)
DT   22-FEB-1995  (first entry)
XX
DE   GAP protein Ira2.
XX
KW   Ras; GTPase activating protein; GAP; GAP related domain; GRD; RAS2;
KW   v-Ras; heat shock; neurofibromatosis type 1; NF1.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO9416069-A2.
XX
PD   21-JUL-1994.
XX
PF   12-JAN-1994;   94WO-US000198.
XX
PR   15-JAN-1993;   93US-00004824.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Nakafuku M,  Kaziro Y;
XX
DR   WPI; 1994-249216/30.
XX
PT   Blocking Ras-induced effects on a cell - by introducing a GTPase
PT   activating protein to the cell, used esp. in treatment of cancers.
XX
PS   Disclosure; Page 63-72; 87pp; English.
XX
CC   Human neurofibromatosis type 1 (NF1)-GAP related domain (GRD) mutant
CC   clones NF201 (given in AAR59221) and NF204 (AAR59922) show strong
CC   suppression activity for RAS2Val19, and inhibit v-Ras-induced
CC   transformation in mammalian cells. The mutation sites of these proteins
CC   were located in one of the most conserved regions of GRD. These sites
CC   were compared with those of other GRD family proteins, yeast Ira2
CC   (AAR59926) and Ira1 (AAR59923), human GAP (AAR59924) and
CC   Schizosaccharomyces pombe Gap1 (AAR59925). (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 3079 AA;

  Query Match             77.3%;  Score 34;  DB 2;  Length 3079;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTKRGL 8
              |||||||
Db       3029 RLTKRGL 3035


RESULT 11
AAY30689
ID   AAY30689 standard; peptide; 10 AA.
XX
AC   AAY30689;
XX
DT   17-NOV-1999  (first entry)
XX
DE   Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW   Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW   low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946598-A1.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;   99WO-US004805.
XX
PR   10-MAR-1998;   98US-0077618P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Innerarity TL,  Boren JOS;
XX
DR   WPI; 1999-551509/46.
XX
PT   Identifying compounds which affect binding of low density lipoprotein
PT   with proteoglycan, used for, e.g. obtaining compounds for reducing
PT   atherosclerosis.
XX
PS   Claim 17; Page 57; 70pp; English.
XX
CC   AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC   receptor mutations. They were created to identify compounds which
CC   modulate atherosclerosis. The peptides are derived from amino acids 3358
CC   to 3367 of apoB100. The method comprises detecting compounds which affect
CC   low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC   can be used for identifying compounds which disrupt LDL-PG binding
CC   without inhibiting LDL receptor binding. Such compounds can be used to
CC   reduce or prevent the formation of atherosclerotic lesions and prevent
CC   atherosclerosis. The transgenic non-human animals and mammals which
CC   express human apo-B100 can be used as an in vivo model system for the
CC   study of atherosclerosis, and in vivo assay methods for identifying
CC   compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC   also be used to identify compounds which result in an increase in
CC   atherosclerotic regions. Thus the assays may be used to determine whether
CC   a particular food or drug composition tends to stimulate or inhibit the
CC   formation of atherosclerotic lesions. The polynucleotides can also be
CC   used in gene therapy for preventing or reducing the severity of
CC   atherosclerosis in an animal or mammal
XX
SQ   Sequence 10 AA;
```

```
  Query Match             76.1%;  Score 33.5;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 3.6;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 TRLT-KRGLK 9
              |||| |||||
Db          1 TRLTDKRGLK 10
```

RESULT 12
AAY30688
ID AAY30688 standard; peptide; 10 AA.
XX
AC AAY30688;
XX
DT 17-NOV-1999 (first entry)
XX
DE Apo-B100 derived peptide showing a proteoglycan receptor mutation.
XX
KW Apo-B100; proteoglycan receptor mutation; atherosclerosis;
KW low density lipoprotein; proteoglycan; LDL; atherosclerotic lesion.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946598-A1.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004805.
XX
PR 10-MAR-1998; 98US-0077618P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Innerarity TL, Boren JOS;
XX
DR WPI; 1999-551509/46.
XX
PT Identifying compounds which affect binding of low density lipoprotein
PT with proteoglycan, used for, e.g. obtaining compounds for reducing
PT atherosclerosis.
XX
PS Claim 17; Page 57; 70pp; English.
XX
CC AAY30582-Y30700 represent apo-B100 derived peptides showing proteoglycan
CC receptor mutations. They were created to identify compounds which
CC modulate atherosclerosis. The peptides are derived from amino acids 3358
CC to 3367 of apoB100. The method comprises detecting compounds which affect
CC low density lipoprotein (LDL) binding with proteoglycan (PG). The method
CC can be used for identifying compounds which disrupt LDL-PG binding
CC without inhibiting LDL receptor binding. Such compounds can be used to
CC reduce or prevent the formation of atherosclerotic lesions and prevent
CC atherosclerosis. The transgenic non-human animals and mammals which
CC express human apo-B100 can be used as an in vivo model system for the
CC study of atherosclerosis, and in vivo assay methods for identifying
CC compounds which modulate atherosclerosis and/or LDL-PG binding. They can
CC also be used to identify compounds which result in an increase in
CC atherosclerotic regions. Thus the assays may be used to determine whether
CC a particular food or drug composition tends to stimulate or inhibit the
CC formation of atherosclerotic lesions. The polynucleotides can also be
CC used in gene therapy for preventing or reducing the severity of
CC atherosclerosis in an animal or mammal
XX
SQ Sequence 10 AA;

```
  Query Match             76.1%;  Score 33.5;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 3.6;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 TRLT-KRGLK 9
              |||| |||||
Db          1 TRLTEKRGLK 10
```

RESULT 13
AAW57205
ID AAW57205 standard; peptide; 11 AA.
XX
AC AAW57205;
XX
DT 03-AUG-1998 (first entry)
XX
DE Apo B binding site peptide 2.
XX
KW Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW growth supplement; non-natural lipid particle; low density lipoprotein;
KW LDL; receptor component; apo B100 receptor site.
XX
OS Synthetic.
XX
PN WO9813385-A2.
XX
PD 02-APR-1998.
XX
PF 25-SEP-1997; 97WO-GB002610.
XX
PR 27-SEP-1996; 96GB-00020153.
XX
PA (UYST) UNIV STRATHCLYDE.
XX
PI Halbert GW, Owens MD, Baillie G;
XX
DR WPI; 1998-230637/20.
XX
PT Non-natural lipid particle comprising peptide binding to apo B protein
PT receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT that express this receptor.
XX
PS Claim 12; Page 52; 73pp; English.
XX
CC The present sequence represents a specifically claimed Apo B binding site
CC peptide which can be used as a component of a non-naturally occurring,
CC receptor-competent low density lipoprotein (LDL) particle of the present
CC invention. The LDL particle comprises at least 1 peptide component that
CC has at least 1 binding site for an apo B protein receptor and at least 1
CC lipophilic substituent. Also described in the invention are peptides
CC containing an apo B binding sequence with at least 70% identity with
CC sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their dimers. Non-
CC naturally occurring, receptor-competent LDL particles are useful as: (i)
CC drug-targeting vectors for delivering anticancer drugs to cancer cells
CC that express an apo B protein receptor, and (ii) additives for cell
CC culture media especially as growth supplements. Non-naturally occurring,
CC receptor-competent LDL particles do not require the complete apo B
CC sequence, which is large and tends to aggregate, to provide binding
CC affinity to an apo B protein receptor
XX
SQ Sequence 11 AA;

```
  Query Match             76.1%;  Score 33.5;  DB 2;  Length 11;
  Best Local Similarity   90.0%;  Pred. No. 4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 TRLT-KRGLK 9
              |||| |||||
Db          2 TRLTRKRGLK 11
```

RESULT 14
AAW57207
ID AAW57207 standard; peptide; 13 AA.
XX
AC AAW57207;
XX
DT 03-AUG-1998 (first entry)

```
XX
DE   Apo B 100 binding site peptide analogue peptide B.
XX
KW   Apo B; binding site; receptor; cancer; drug delivery; anticancer;
KW   growth supplement; non-natural lipid particle; low density lipoprotein;
KW   LDL; receptor component; apo B100 receptor site.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "attached to retinoic acid"
XX
PN   WO9813385-A2.
XX
PD   02-APR-1998.
XX
PF   25-SEP-1997;   97WO-GB002610.
XX
PR   27-SEP-1996;   96GB-00020153.
XX
PA   (UYST ) UNIV STRATHCLYDE.
XX
PI   Halbert GW,  Owens MD,  Baillie G;
XX
DR   WPI; 1998-230637/20.
XX
PT   Non-natural lipid particle comprising peptide binding to apo B protein
PT   receptor - useful as, e.g. vector for delivering drugs to cancer cells
PT   that express this receptor.
XX
PS   Claim 13; Fig 7; 73pp; English.
XX
CC   The present sequence represents a specifically claimed Apo B 100 binding
CC   site peptide analogue which can be used as a component of a non-
CC   naturally occurring, receptor-competent low density lipoprotein (LDL)
CC   particle of the present invention. The LDL particle comprises at least 1
CC   peptide component that has at least 1 binding site for an apo B protein
CC   receptor and at least 1 lipophilic substituent. Also described in the
CC   invention are peptides containing an apo B binding sequence with at least
CC   70% identity with sequences: KAEYKKNKHRH (1) or TTRLTRKRGLK (2), or their
CC   dimers. Non-naturally occurring, receptor-competent LDL particles are
CC   useful as: (i) drug-targeting vectors for delivering anticancer drugs to
CC   cancer cells that express an apo B protein receptor, and (ii) additives
CC   for cell culture media especially as growth supplements. Non-naturally
CC   occurring, receptor-competent LDL particles do not require the complete
CC   apo B sequence, which is large and tends to aggregate, to provide binding
CC   affinity to an apo B protein receptor
XX
SQ   Sequence 13 AA;

  Query Match             76.1%;  Score 33.5;  DB 2;  Length 13;
  Best Local Similarity   90.0%;  Pred. No. 4.7;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 TRLT-KRGLK 9
              |||| |||||
Db          3 TRLTRKRGLK 12


RESULT 15
AAW41261
ID   AAW41261 standard; peptide; 15 AA.
XX
AC   AAW41261;
XX
DT   19-MAY-1998  (first entry)
XX
DE   Apolipoprotein B-100 fragment.
XX
KW   Anti-coagulant; apolipoprotein B-100; apoB-100; metastatic spread;
KW   thromboplastin-mediated process; cancer; inhibitor; blood coagulation;
KW   angiogenesis; cellular differentiation; apoptosis; KRAD-14;
KW   prothrombinase complex.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9743311-A1.
XX
PD   20-NOV-1997.
XX
PF   09-MAY-1997;   97WO-GB001255.
XX
PR   09-MAY-1996;   96GB-00009702.
XX
PA   (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.
XX
PI   Bruckdorfer KR,  Ettelaie C;
XX
DR   WPI; 1998-008798/01.
XX
PT   Peptide fragments of apo:lipo:protein B-100 with anticoagulant activity -
PT   used for treating or preventing coagulation, inhibiting angiogenesis,
PT   cell differentiation and apoptosis.
XX
PS   Disclosure; Page 22; 60pp; English.
XX
CC   This sequence is an example of the peptide of the invention. It has the
CC   formula (I), or their variants with one or more internal deletions,
CC   insertions or substitutions, while retaining anti-coagulant properties of
CC   apolipoprotein B-100 (apoB-100). Z1-KAQ-X1-KKNKHRHS-X2-T-Z2 (I) X1 = S or
CC   Y; X2 = T or I; Z1 = the N terminus of the peptide, or 1-47 amino acids
CC   (aa); Z2 = the C terminus of the peptide, a terminal amide group or 1-77
CC   aa. Compositions containing the peptide are used for simultaneous,
CC   separate or sequential treatment of cancer, particularly to prevent
CC   metastatic spread. They are also used to inhibit thromboplastin-mediated
CC   processes, specifically to prevent or reduce blood coagulation (e.g.
CC   during or after surgery or in cases of heart attack, stroke etc.) and to
CC   inhibit angiogenesis, cellular differentiation or apoptosis. KRAD-14,
CC   which is active as such or as part of a 98-aa peptide, inhibits
CC   activation of the prothrombinase complex; and prevents activation of
CC   factor VII on the surface of thromboplastin and of platelets by thrombin.
CC   It binds to the residues 58-66 of thromboplastin. Since (I) are much
CC   smaller than apoB-100, they act more quickly
XX
SQ   Sequence 15 AA;

  Query Match             76.1%;  Score 33.5;  DB 2;  Length 15;
  Best Local Similarity   90.0%;  Pred. No. 5.5;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 TRLT-KRGLK 9
              |||| |||||
Db          1 TRLTRKRGLK 10


Search completed: December 29, 2004, 12:28:51
Job time : 55.9205 secs
```

THIS PAGE IS BLANK

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:15:57 ; Search time 8.69318 Seconds
(without alignments)
99.613 Million cell updates/sec

Title: US-09-823-418-14
Perfect score: 44
Sequence: 1 TRLTKRGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	36	81.8	427	2	A84155	hypothetical prote
2	35	79.5	631	2	T29926	hypothetical prote
3	34	77.3	99	2	A87912	protein B0205.5 [i
4	34	77.3	248	2	S77172	glucose dehydrogen
5	34	77.3	840	2	F82937	DNA topoisomerase
6	34	77.3	3079	1	RGBYI2	probable GTPase-ac
7	33.5	76.1	596	2	S32802	apolipoprotein B -
8	33.5	76.1	4563	1	LPHUB	apolipoprotein B-1
9	33	75.0	147	2	AI1315	nucleoside diphosp
10	33	75.0	151	1	R3KW13	ribosomal protein
11	33	75.0	470	2	S39733	amino acid permeas
12	33	75.0	633	2	T05005	hypothetical prote
13	33	75.0	800	2	S51368	ribosomal protein
14	32	72.7	114	1	ETHUL	lymphotactin precu
15	32	72.7	257	2	I40170	hypothetical prote
16	32	72.7	263	2	G70179	spermidine/putresc
17	32	72.7	271	2	S27422	peroxisomal assemb
18	32	72.7	275	2	A72253	lytB protein - The
19	32	72.7	306	2	T50120	hypothetical prote
20	32	72.7	345	1	JH0185	D-amino-acid oxida
21	32	72.7	347	1	OXPGDA	D-amino-acid oxida
22	32	72.7	347	1	S01340	D-amino-acid oxida
23	32	72.7	347	1	JX0132	D-amino-acid oxida
24	32	72.7	376	2	A98030	coproporphyrinogen
25	32	72.7	376	2	B95164	hypothetical prote
26	32	72.7	389	2	E96835	unknown protein F5
27	32	72.7	393	2	H75444	branched-chain ami
28	32	72.7	454	2	AG1977	hypothetical prote
29	32	72.7	625	2	B86875	metal transporting
30	32	72.7	871	2	T07863	probable polyprote
31	32	72.7	1145	2	A59251	myosin - Acetabula
32	31	70.5	85	2	G81430	hypothetical prote
33	31	70.5	114	1	ETMSL	lymphotactin precu
34	31	70.5	180	2	C72769	probable diphtheri
35	31	70.5	232	2	T29841	hypothetical prote
36	31	70.5	364	2	C87455	alanine racemase [
37	31	70.5	378	2	AI2180	hypothetical prote
38	31	70.5	462	2	T17948	ABC transporter pr
39	31	70.5	474	2	T44424	dihydrolipoamide d
40	31	70.5	482	2	S18210	hypothetical prote
41	31	70.5	485	2	T25684	hypothetical prote
42	31	70.5	559	2	T42998	Ras-binding protei
43	31	70.5	572	2	T30947	hypothetical prote
44	31	70.5	605	2	T38932	probable sulfur me
45	31	70.5	613	2	A88684	protein AC7.2 [imp

ALIGNMENTS

RESULT 1
A84155
hypothetical protein BH4041 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A84155
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A84155
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <STO>
A;Cross-references: UNIPROT:Q9K5P7; GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH4041

Query Match 81.8%; Score 36; DB 2; Length 427;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347
```

RESULT 2
T29926
hypothetical protein T03G11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29926
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid T03G11.
A;Reference number: Z20709
A;Accession: T29926
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-631 <GEI>
A;Cross-references: UNIPROT:Q22126; EMBL:U41272; PIDN:AAA82452.1; CESP:T03G11.1
C;Genetics:
A;Gene: CESP:T03G11.1
A;Introns: 107/3; 214/3; 250/1; 306/2; 364/3; 405/2; 451/3; 522/1; 576/2

Query Match 79.5%; Score 35; DB 2; Length 631;
Best Local Similarity 87.5%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLTKRGLK 9
```

```
              ||||:|||
Db        539 RLTKQGLK 546


RESULT 3
A87912
protein B0205.5 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: A87912
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: A87912
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <STO>
A;Cross-references: UNIPROT:O61740; GB:chr_I; PIDN:AAC16987.1; PID:g3150469; GSPDB:GN000
C;Genetics:
A;Gene: B0205.5
A;Map position: 1

  Query Match             77.3%;  Score 34;  DB 2;  Length 99;
  Best Local Similarity   66.7%;  Pred. No. 9.8;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ::|||||:|
Db         65 SKLTKRGVK 73


RESULT 4
S77172
glucose dehydrogenase - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sll1709
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S77172
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S77172
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-248 <KAN>
A;Cross-references: UNIPROT:P73684; EMBL:D90908; GB:AB001339; NID:g1652725; PIDN:BAA1773
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: gdh
C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F;6-186/Domain: short-chain alcohol dehydrogenase homology <SADH>

  Query Match             77.3%;  Score 34;  DB 2;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLTKRGL 8
              |||||||
Db        157 RLTKRGL 163


RESULT 5
F82937
DNA topoisomerase chain A, DNA gyrase A subunit UU082 [imported] - Ureaplasma urealyticu
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: F82937
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: F82937
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-840 <GLA>
A;Cross-references: GB:AE002107; GB:AF222894; NID:g6899022; PIDN:AAF30487.1; GSPDB:GN001
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: gyrA; UU082
A;Genetic code: SGC3
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (

  Query Match             77.3%;  Score 34;  DB 2;  Length 840;
  Best Local Similarity   87.5%;  Pred. No. 73;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTKRGLK 9
              |||||| |
Db        744 RLTKRGAK 751


RESULT 6
RGBYI2
probable GTPase-activating protein IRA2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: GLC4 protein; protein O0985; protein YOL081w; protein YOL0951
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1991 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: S66775; S66774; A35656; S48254; S50426; S11190; S38505
R;Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66775
A;Accession: S66775
A;Molecule type: DNA
A;Residues: 1-2423 <ZUM>
A;Cross-references: UNIPROT:P19158; UNIPROT:O13592; EMBL:Z74823; GSPDB:GN00015; MIPS:YOL
A;Experimental source: strain S288C
R;Alexandraki, D.; Katsoulou, C.; Tzermia, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66756
A;Accession: S66774
A;Molecule type: DNA
A;Residues: 1983-3079 <ALE>
A;Cross-references: EMBL:Z74823; GSPDB:GN00015; MIPS:YOL081w
A;Experimental source: strain S288C
R;Tanaka, K.; Nakafuku, M.; Tamanoi, F.; Kaziro, Y.; Matsumoto, K.; Toh-e, A.
Mol. Cell. Biol. 10, 4303-4313, 1990
A;Title: IRA2, a second gene of Saccharomyces cerevisiae that encodes a protein with a d
A;Reference number: A35656; MUID:90318397; PMID:2164637
A;Accession: A35656
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-2308,'V',2310-2316,'I',2318-3079 <TAN>
A;Cross-references: EMBL:M33779; NID:g171761; PIDN:AAA34710.1; PID:g171762
R;Zumstein, E.; Griffin, H.; Schweizer, M.
Yeast 10, 1383-1387, 1994
A;Title: Sequence of a 10.27 kb segment on the left arm of chromosome XV from Saccharomy
A;Reference number: S48253; MUID:95208358; PMID:7900427
A;Accession: S48254
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2423 <ZU2>
A;Cross-references: EMBL:X75449; NID:g414079; PIDN:CAA53202.1; PID:g414081
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993
R;Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
submitted to the EMBL Data Library, December 1994
A;Reference number: S50410
A;Accession: S50426
A;Molecule type: DNA
A;Residues: 1-2423 <ZU3>
```

A;Cross-references: EMBL:X83121; NID:g600461; PIDN:CAA58201.1; PID:g600480
C;Genetics:
A;Gene: SGD:IRA2; MIPS:YOL081w
A;Cross-references: SGD:S0005441; MIPS:YOL081w
A;Map position: 15L
C;Superfamily: regulatory protein IRA2; ras-specific GAP catalytic domain homology
C;Keywords: transmembrane protein
F;693-709/Domain: transmembrane #status predicted <TM1>
F;1135-1151/Domain: transmembrane #status predicted <TM2>
F;1701-1910/Domain: ras-specific GAP catalytic domain homology <GAP>
F;1842-1858/Domain: transmembrane #status predicted <TM3>
F;2318-2334/Domain: transmembrane #status predicted <TM4>
F;2562-2578/Domain: transmembrane #status predicted <TM5>

Query Match 77.3%; Score 34; DB 1; Length 3079;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTKRGL 8
||||||| |
Db 3029 RLTKRGL 3035

RESULT 7
S32802
apolipoprotein B - crab-eating macaque (fragment)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S32802
R;Pape, M.E.; Castle, C.K.; Murray, R.W.; Funk, G.M.; Hunt, C.E.; Marotti, K.R.; Melchic
Biochim. Biophys. Acta 1086, 326-334, 1991
A;Title: Apo B metabolism in the cynomolgus monkey: evidence for post-transcriptional re
A;Reference number: S32802; MUID:92075708; PMID:1742325
A;Accession: S32802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-596 <PAP>
A;Cross-references: UNIPROT:Q28473; EMBL:X15737; NID:g38047; PIDN:CAA33755.1; PID:g93012
C;Superfamily: apolipoprotein B

Query Match 76.1%; Score 33.5; DB 2; Length 596;
Best Local Similarity 90.0%; Pred. No. 67;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLT-KRGLK 9
|||| ||||| |
Db 226 TRLTRKRGLK 235

RESULT 8
LPHUB
apolipoprotein B-100 precursor - human
N;Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A27850; A25679; A25263; A25267; A25266; A24320; A24684; A23817; A25774; A26
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058
R;Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Scc
DNA 6, 363-372, 1987
A;Title: DNA sequence of the human apolipoprotein B gene.
A;Reference number: A27850; MUID:88003974; PMID:3652907
A;Accession: A27850
A;Molecule type: DNA
A;Residues: 1-617,'A',619-1929,'F',1931-3318,'D',3320-3426,'T',3428-3431,'Q',3433-3731,'
A;Cross-references: UNIPROT:P04114; UNIPROT:P78482; UNIPROT:P78479; UNIPROT:Q9UMN0; UNIP
R;Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I.
EMBO J. 5, 3495-3507, 1986
A;Title: The complete sequence and structural analysis of human apolipoprotein B-100: re
A;Reference number: A91058; MUID:87161758; PMID:3030729
A;Accession: A25679
A;Molecule type: mRNA
A;Residues: 1-11,15-2539,'S',2541-3823,'R',3825-4563 <CLA>
A;Note: 1109-Asp was also found

R;Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McC
Nucleic Acids Res. 14, 7501-7503, 1986
A;Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A;Reference number: A93639; MUID:87016385; PMID:3763409
A;Accession: A25263
A;Molecule type: mRNA
A;Residues: 1-272,'N',274-617,'A',619-1217,'E',1219-2091,'V',2093-2364,'T',2366-2679,'Q
A;Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R;Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer J
Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A;Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A;Reference number: A94134; MUID:87041416; PMID:3464946
A;Accession: A25267
A;Molecule type: mRNA
A;Residues: 1-617,'A',619-703,'P',705-792,'R',794-1270,'S',1272-1866,'G',1868-2036,'N',
4189-4220,'M',4222-4563 <LAW>
A;Note: the codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R;Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.
J. Biol. Chem. 261, 12918-12921, 1986
A;Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A;Reference number: A92556; MUID:87008488; PMID:3759943
A;Accession: A25266
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-328,'V',330-644,'I',646-918,'P',920-3318,'D',3320-3426,'T',3428
9-4132,'G',4134-4180,'E',4182-4563 <CHE>
A;Cross-references: GB:J02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A;Note: a total of 2366 residues were confirmed by direct sequencing of tryptic peptide
R;Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.;
Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A;Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotei
A;Reference number: A24320; MUID:86287319; PMID:3461454
A;Accession: A24320
A;Molecule type: mRNA
A;Residues: 1-97,'I',99-617,'A',619-941,'YYIWSLPPKP',951-1138,'PTGRLPNCFSNGLICYSLWLHSFQ
A;Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R;Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor,
Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A;Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A;Reference number: A24684; MUID:86094221; PMID:3001697
A;Accession: A24684
A;Molecule type: mRNA
A;Residues: 485-617,'A',619-1044 <LA2>
A;Cross-references: GB:M12480; NID:g178791; PIDN:AAA51751.1; PID:g178792
R;Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; K
Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A;Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipop
A;Reference number: A94088; MUID:86149325; PMID:3513177
A;Accession: A23817
A;Molecule type: mRNA
A;Residues: 1-291 <PRO>
A;Cross-references: GB:M12681; NID:g178797; PIDN:AAA51753.1; PID:g178798
R;Deeb, S.S.; Motulsky, A.G.; Albers, J.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A;Title: A partial cDNA clone for human apoliprotein B.
A;Reference number: A25774; MUID:85270450; PMID:3860836
A;Accession: A25774
A;Molecule type: mRNA
A;Residues: 709-791,'SSSWKAASHGCPHSAGD',810-906 <DEE>
A;Cross-references: GB:K03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R;Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G.
Gene 49, 29-51, 1986
A;Title: Analysis of the human apolipoprotein B gene; complete structure of the B-74 re
A;Reference number: A91565; MUID:87191999; PMID:2883086
A;Accession: A26533
A;Molecule type: mRNA
A;Residues: 1282-2721,2742-3290,'L',3292-3336,'N',3338-3948,'F',3950-3963,'Y',3965-4180
A;Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R;Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yaman
Biochemistry 26, 5478-5486, 1987
A;Title: Structural comparison of human apolipoproteins B-48 and B-100.
A;Reference number: A29671; MUID:88050832; PMID:3676265
A;Accession: A29671
A;Molecule type: mRNA

A;Residues: 1671-2323,'PYW',2327-2352,'H',2354-2398 <HAR>
A;Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732
R;Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F.E.;
Atherosclerosis 58, 277-289, 1985
A;Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than or
A;Reference number: A90084; MUID:86130855; PMID:3841481
A;Accession: A29287
A;Molecule type: mRNA
A;Residues: 3846-4298 <SHO>
R;Pfitzner, R.; Wagener, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A;Title: Isolation, expression and characterization of a human apolipoprotein B 100-spec
A;Reference number: A25572; MUID:87076044; PMID:3024665
A;Accession: A25572
A;Molecule type: mRNA
A;Residues: 4219-4337,'S',4339-4563 <PFI>
A;Cross-references: GB:M36676
R;Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J.T.;
Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A;Reference number: A24738; MUID:86042646; PMID:2932736
A;Accession: A24738
A;Molecule type: mRNA
A;Residues: 'N',3729-3731,'I',3733-3875,'A',3877-3948,'F',3950-3963,'Y',3965-3982,'S',39
A;Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R;Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Cai
Science 238, 363-366, 1987
A;Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in
A;Reference number: A40133; MUID:88018019; PMID:3659919
A;Accession: B40133
A;Molecule type: mRNA
A;Residues: 2165-2179 <CH1>
A;Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A;Note: this mRNA includes the stop codon of the organ-specific mRNA for apo48
A;Accession: A40133
A;Molecule type: protein
A;Residues: 51-75;101-110;129-139;158-174;197-207;276-287;298-304;306-314;526-532;538-55
36;1486-1498;1537-1556;1563-1572;1601-1610;1647-1661;1697-1724;1770-1781;1859-1897;1968-
A;Note: these fragments were derived from apo48
R;Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P.
Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A;Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanism p
A;Reference number: A28002; MUID:88106542; PMID:3426612
A;Accession: A28002
A;Molecule type: mRNA
A;Residues: 2129-2179,2181-2235 <HA2>
A;Cross-references: GB:M18471
A;Experimental source: intestine
A;Note: this mRNA from intestine includes a stop codon created by RNA editing in place c
R;Mehrabian, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner, T
Nucleic Acids Res. 13, 6937-6953, 1985
A;Title: Human apolipoprotein B: identification of cDNA clones and characterization of m
A;Reference number: A24269; MUID:86041888; PMID:3903660
A;Accession: A24269
A;Molecule type: mRNA
A;Residues: 3056-3159 <MEH>
A;Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R;Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.
Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A;Title: Identification of a novel in-frame translational stop codon in human intestine
A;Reference number: A29659; MUID:88049670; PMID:2445342
A;Accession: A29659
A;Molecule type: mRNA
A;Residues: 2169-2179 <HOS>
A;Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A;Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intest
ch encodes the 250K apoB-48, CAA encoding 2180-Gln is substituted by the stop codon TAA,
R;Yang, C.; Kim, T.W.; Weng, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A;Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A;Reference number: A35783; MUID:90319144; PMID:2115173
A;Contents: disulfide bonds
A;Accession: A35783
A;Molecule type: protein
A;Residues: 28-41;76-97,'I',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5

A;Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R;LeBoeuf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A;Title: Human apolipoprotein B: partial amino acid sequence.
A;Reference number: A22006; MUID:84208786; PMID:6373369
A;Accession: A22006
A;Molecule type: protein
A;Residues: 873-892,'K',894-896 <LE1>
A;Accession: B22006
A;Molecule type: protein
A;Residues: 3113,'L',3115-3130,'R',3132-3133,'P',3135-3136,'R' <LE2>
R;Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C ;
J. Biol. Chem. 261, 15364-15367, 1986
A;Title: Structure of the human apolipoprotein B gene.
A;Reference number: A92564; MUID:87057153; PMID:2946672
A;Contents: annotation; gene structure
R;Wagener, R.; Pfitzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A;Title: Studies on the organization of the human apolipoprotein B 100 gene.
A;Reference number: A90715; MUID:87271140; PMID:2886136
A;Contents: annotation; gene structure
R;Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A;Title: Human apolipoprotein B-100 heparin-binding sites.
A;Reference number: A92605; MUID:87280197; PMID:3301850
A;Contents: annotation; heparin binding and disulfide bond
R;Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A;Title: Apolipoprotein B is a calcium binding protein.
A;Reference number: A90125; MUID:86242245; PMID:3087360
A;Contents: annotation; calcium binding
R;Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A;Title: Molecular cloning of human apolipoprotein B cDNA.
A;Reference number: I37178; MUID:86093680; PMID:3841204
A;Accession: I37180

Query Match 76.1%; Score 33.5; DB 1; Length 4563;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 TRLT-KRGLK 9
|||| |||||
Db 3385 TRLTRKRGLK 3394

RESULT 9
AI1315
nucleoside diphosphate kinase homolog ndk [imported] - Listeria monocytogenes (strain EC
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C;Accession: AI1315
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1315
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-147 <GLA>
A;Cross-references: UNIPROT:Q8Y5X4; GB:NC_003210; PIDN:CAD00007.1; PID:g16411382; GSPDB
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: ndk
C;Superfamily: Nucleoside diphosphate kinase

Query Match 75.0%; Score 33; DB 2; Length 147;

```
  Best Local Similarity   66.7%;  Pred. No. 23;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||: |:|||
Db         23 TRIEKKGLK 31


RESULT 10
R3KW13
ribosomal protein S13.e, cytosolic - nematode (Brugia pahangi)
N;Alternate names: 17.4K protein
C;Species: Brugia pahangi
C;Date: 30-Sep-1991 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C;Accession: S32687; S14440; S06771
R;Ellenberger, D.L.; Pieniazek, N.J.; Lammie, P.J.
submitted to the EMBL Data Library, January 1992
A;Description: Developmental modulation of relative gene numbers in a parasitic nematode
A;Reference number: S32687
A;Accession: S32687
A;Molecule type: DNA
A;Residues: 1-151 <ELL>
A;Cross-references: UNIPROT:Q17274; EMBL:X63714; NID:g297070; PIDN:CAA45247.1; PID:g2970
R;Ellenberger, D.L.; Pieniazek, N.J.; Lammie, P.J.
Nucleic Acids Res. 17, 10121, 1989
A;Title: Nucleotide sequence of Brugia pahangi 17.4 kD protein.
A;Reference number: S14440; MUID:90098795; PMID:2602125
A;Accession: S14440
A;Molecule type: mRNA
A;Residues: 1-26,'K',28-36,'V',38-106,'Q',108-118,'Q',120-123,'R',125-151 <EL2>
A;Cross-references: EMBL:X16591
C;Genetics:
A;Introns: 43/3; 78/3; 107/3
C;Superfamily: rat ribosomal protein S13; eubacterial ribosomal protein S15 homology
C;Keywords: protein biosynthesis; ribosome
F;2-151/Product: ribosomal protein S13.e #status predicted <MAT>
F;82-148/Domain: eubacterial ribosomal protein S15 homology <ES15>

  Query Match             75.0%;  Score 33;  DB 1;  Length 151;
  Best Local Similarity   66.7%;  Pred. No. 23;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||| |:||:
Db         38 TRLAKKGLR 46


RESULT 11
S39733
amino acid permease rocC - Bacillus subtilis
N;Alternate names: protein ipa-77d
C;Species: Bacillus subtilis
C;Date: 07-Oct-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S39733; G69693
R;Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
, A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. 10, 371-384, 1993
A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
A;Reference number: S39655; MUID:95020537; PMID:7934828
A;Accession: S39733
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-470 <GLA>
A;Cross-references: UNIPROT:P39636; EMBL:X73124; NID:g413923; PIDN:CAA51634.1; PID:g4140
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G69693
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-470 <KUN>
A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15803.1; PID:g2636312
A;Experimental source: strain 168
C;Genetics:
A;Gene: rocC
C;Superfamily: arginine permease
C;Keywords: amino acid transport; glycoprotein; transmembrane protein

  Query Match             75.0%;  Score 33;  DB 2;  Length 470;
  Best Local Similarity   75.0%;  Pred. No. 68;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGL 8
              |||||:|:
Db        327 TRLTKKGV 334


RESULT 12
T05005
hypothetical protein T19P19.70 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05005
R;Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15394
A;Accession: T05005
A;Molecule type: DNA
A;Residues: 1-633 <BEV>
A;Cross-references: UNIPROT:O65655; EMBL:AL022605
A;Experimental source: cultivar Columbia; BAC clone T19P19
C;Genetics:
A;Map position: 4
A;Introns: 385/1; 448/1; 498/3
A;Note: T19P19.70

  Query Match             75.0%;  Score 33;  DB 2;  Length 633;
  Best Local Similarity   87.5%;  Pred. No. 90;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLTKRGLK 9
              ||| ||||
Db         30 RLTTRGLK 37


RESULT 13
S51368
ribosomal protein S3 - Chlamydomonas eugametos chloroplast
C;Species: chloroplast Chlamydomonas eugametos
C;Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S51368
R;Turmel, M.; Otis, C.
Curr. Genet. 27, 54-61, 1994
A;Title: The chloroplast gene cluster containing psbF, psbL, petG and rps3 is conserved
A;Reference number: S51365; MUID:95269309; PMID:7750147
A;Accession: S51368
A;Molecule type: DNA
A;Residues: 1-800 <TUR>
A;Cross-references: UNIPROT:P46307; EMBL:L29282; NID:g575472; PID:g575476
C;Genetics:
A;Gene: rps3
A;Genome: chloroplast
```

C;Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 75.0%; Score 33; DB 2; Length 800;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TRLTKRGLK 9
:|:|||| ||
Db 266 TKLTKRQLK 274

RESULT 14
ETHUL
lymphotactin precursor - human
N;Alternate names: activation-induced chemokine-related protein (ATAC); single cysteine
N;Contains: eosinophilotactic peptide
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1981 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C;Accession: S60650; I38978; A03190; I53506
R;Mueller, S.; Dorner, B.; Korthaeuer, U.; Mages, H.W.; D'Apuzzo, M.; Senger, G.; Krocze
Eur. J. Immunol. 25, 1744-1748, 1995
A;Title: Cloning of ATAC, an activation-induced, chemokine-related molecule exclusively
A;Reference number: S60650; MUID:95339892; PMID:7615002
A;Accession: S60650
A;Molecule type: mRNA
A;Residues: 1-114 <MUE>
A;Cross-references: UNIPROT:P47992; EMBL:X86474; NID:g895846; PIDN:CAA60198.1; PID:g8958
R;Kennedy, J.; Kelner, G.S.; Kleyensteuber, S.; Schall, T.J.; Weiss, M.C.; Yssel, H.; Sc
J. Immunol. 155, 203-209, 1995
A;Title: Molecular cloning and functional characterization of human lymphotactin.
A;Reference number: I38978; MUID:95325590; PMID:7602097
A;Accession: I38978
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-114 <KEN>
A;Cross-references: EMBL:U23772; NID:g902001; PIDN:AAC50164.1; PID:g902002
R;Goetzl, E.J.; Austen, K.F.
Proc. Natl. Acad. Sci. U.S.A. 72, 4123-4127, 1975
A;Title: Purification and synthesis of eosinophilotactic tetrapeptides of human lung tis
A;Reference number: A03190; MUID:76078412; PMID:1060093
A;Accession: A03190
A;Molecule type: protein
A;Residues: 22-25 <GOE>
A;Note: 22-Ala was also seen
R;Yoshida, T.; Imai, T.; Kakizaki, M.; Nishimura, M.; Yoshie, O.
FEBS Lett. 360, 155-159, 1995
A;Title: Molecular cloning of a novel C or gamma type chemokine, SCM-1.
A;Reference number: I53506; MUID:95180438; PMID:7875320
A;Accession: I53506
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-114 <YOS>
A;Cross-references: GB:D43768; NID:g927650; PIDN:BAA07825.1; PID:g927651
C;Comment: Lymphotactin is produced by activated T-cells and is chemotactic for some lym
C;Comment: Eosinophilotactic peptide is released from mast cells in lung and other tissu
ially affecting eosinophils, include chemotaxis, chemotactic deactivation, release of en
C;Comment: It has not yet been shown that the previously detected eosinophilotactic pept
C;Genetics:
A;Gene: GDB:SCYC1; LTN; LPTN; ATAC
A;Cross-references: GDB:682094
A;Map position: 1q23-1q25
C;Superfamily: lymphotactin
C;Keywords: chemotaxis; cytokine; lymphokine; mast cell; T-cell
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-21/Domain: propeptide #status predicted <PRO>
F;22-114/Product: lymphotactin #status predicted <MAT>
F;22-25/Product: eosinophilotactic peptide #status predicted <EOP>
F;32-69/Disulfide bonds: #status predicted

Query Match 72.7%; Score 32; DB 1; Length 114;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LTKRGLK 9
:||||||
Db 61 ITKRGLK 67

RESULT 15
I40170
hypothetical protein 2 - Bacillus caldolyticus
C;Species: Bacillus caldolyticus
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Jul-2004
C;Accession: I40170; S34322
R;Ghim, S.Y.; Neuhard, J.
J. Bacteriol. 176, 3698-3707, 1994
A;Title: The pyrimidine biosynthesis operon of the thermophile Bacillus caldolyticus inc
A;Reference number: I40166; MUID:94266723; PMID:8206848
A;Accession: I40170
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-257 <RES>
A;Cross-references: UNIPROT:P46536; EMBL:X73308; NID:g312439; PIDN:CAA51740.1; PID:g3124
C;Superfamily: cytochrome-c3 hydrogenase gamma chain

Query Match 72.7%; Score 32; DB 2; Length 257;
Best Local Similarity 75.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLTKRGLK 9
:|||||:|
Db 129 QLTKRGVK 136

Search completed: December 29, 2004, 12:39:10
Job time : 10.6932 secs

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:13:11 ; Search time 52.5682 Seconds
(without alignments)
98.508 Million cell updates/sec

Title: US-09-823-418-14
Perfect score: 44
Sequence: 1 TRLTKRGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	90.9	423	2	Q7WSQ9	Q7wsq9 arthrobacte
2	38	86.4	188	2	Q8QHI6	Q8qhi6 gallus gall
3	38	86.4	400	1	FXLE_MOUSE	Q8bid8 mus musculu
4	38	86.4	418	1	FXLE_HUMAN	Q8n1e6 homo sapien
5	36	81.8	423	2	Q75TW9	Q75tw9 bacillus ha
6	36	81.8	423	2	BAD18188	Bad18188 bacillus
7	36	81.8	427	2	Q75TN6	Q75tn6 bacillus fi
8	36	81.8	427	2	Q75TQ1	Q75tq1 bacillus al
9	36	81.8	427	2	Q75TX1	Q75tx1 bacillus ha
10	36	81.8	427	2	Q75TX3	Q75tx3 bacillus ha
11	36	81.8	427	2	Q75TX6	Q75tx6 bacillus ha
12	36	81.8	427	2	Q75TX7	Q75tx7 bacillus ha
13	36	81.8	427	2	Q75TZ5	Q75tz5 bacillus ha
14	36	81.8	427	2	Q75TZ7	Q75tz7 bacillus ha
15	36	81.8	427	2	Q75TZ8	Q75tz8 bacillus ha
16	36	81.8	427	2	Q9K5P7	Q9k5p7 bacillus ha
17	36	81.8	427	2	BAD18156	Bad18156 bacillus
18	36	81.8	427	2	BAD18157	Bad18157 bacillus
19	36	81.8	427	2	BAD18158	Bad18158 bacillus
20	36	81.8	427	2	BAD18159	Bad18159 bacillus
21	36	81.8	427	2	BAD18160	Bad18160 bacillus
22	36	81.8	427	2	BAD18161	Bad18161 bacillus
23	36	81.8	427	2	BAD18162	Bad18162 bacillus
24	36	81.8	427	2	BAD18180	Bad18180 bacillus
25	36	81.8	427	2	BAD18181	Bad18181 bacillus
26	36	81.8	427	2	BAD18183	Bad18183 bacillus
27	36	81.8	427	2	BAD18184	Bad18184 bacillus
28	36	81.8	427	2	BAD18185	Bad18185 bacillus
29	36	81.8	427	2	BAD18186	Bad18186 bacillus
30	36	81.8	427	2	BAD18256	Bad18256 bacillus
31	36	81.8	427	2	BAD18259	Bad18259 bacillus
32	36	81.8	427	2	BAD18271	Bad18271 bacillus
33	36	81.8	772	2	Q6CLD8	Q6cld8 kluyveromyc
34	35	79.5	386	2	Q8QHI8	Q8qhi8 brachydanio
35	35	79.5	392	2	Q8QHI7	Q8qhi7 brachydanio
36	35	79.5	411	2	Q6TGX5	Q6tgx5 brachydanio
37	35	79.5	411	2	AAQ97814	Aaq97814 brachydan
38	34	77.3	99	2	O61740	O61740 caenorhabdi
39	34	77.3	156	2	Q73HA4	Q73ha4 wolbachia p
40	34	77.3	156	2	AAS14361	Aas14361 wolbachia
41	34	77.3	248	2	P73684	P73684 synechocyst
42	34	77.3	282	2	Q6ME72	Q6me72 parachlamyd
43	34	77.3	282	2	CAF23127	Caf23127 parachlam
44	34	77.3	335	2	Q8PHW2	Q8phw2 xanthomonas
45	34	77.3	840	1	GYRA_UREPA	Q9pr63 ureaplasma

ALIGNMENTS

RESULT 1
Q7WSQ9
ID Q7WSQ9 PRELIMINARY; PRT; 423 AA.
AC Q7WSQ9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transporter protein.
OS Arthrobacter ilicis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=43665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rue61a;
RX MEDLINE=22753791; PubMed=12730200;
RA Parschat K., Hauer B., Kappl R., Kraft R., Huettermann J., Fetzner S.;
RT "Gene Cluster of Arthrobacter ilicis R.61a Involved in the Degradation
RT of Quinaldine to Anthranilate. Characterization and Functional
RT Expression of the Quinaldine 4-oxidase qoxLMS Genes.";
RL J. Biol. Chem. 278:27483-27494(2003).
DR EMBL; AJ537472; CAD61041.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
SQ SEQUENCE 423 AA; 43696 MW; BB11CBADA85DF241 CRC64;

Query Match 90.9%; Score 40; DB 2; Length 423;
Best Local Similarity 88.9%; Pred. No. 8.3;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TRLTKRGLK 9
              |||||:|||
Db        207 TRLTKQGLK 215
```

RESULT 2
Q8QHI6
ID Q8QHI6 PRELIMINARY; PRT; 188 AA.
AC Q8QHI6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PPA (Fragment).
GN Name=PPA;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.

```
RX   MEDLINE=21972450; PubMed=11976951;
RA   Das T., Purkayastha-Mukherjee C., D'Angelo J., Weir M.;
RT   "A conserved F-box gene with unusual transcript localization.";
RL   Dev. Genes Evol. 212:134-140(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Das T.K., Purkayastha-Mukherjee C., D'Angelo J., Weir M.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF467464; AAL75968.1; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007089; LRR_cys.
DR   Pfam; PF00560; LRR; 4.
FT   NON_TER       1      1
FT   NON_TER     188    188
SQ   SEQUENCE   188 AA;  20629 MW;  21702832DA5CE865 CRC64;

  Query Match             86.4%;  Score 38;  DB 2;  Length 188;
  Best Local Similarity   77.8%;  Pred. No. 9.6;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||||:
Db        161 TRITKRGLE 169


RESULT 3
FXLE_MOUSE
ID   FXLE_MOUSE     STANDARD;      PRT;   400 AA.
AC   Q8BID8; Q8R5H7; Q8VDT7; Q922N5;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   F-box/LRR-repeat protein 14 (F-box and leucine-rich repeat protein
DE   14).
GN   Name=Fbxl14; Synonyms=Ppa;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Breast tumor, and Heart;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 207-390 FROM N.A.
RX   MEDLINE=21972450; PubMed=11976951; DOI=10.1007/s00427-002-0222-7;
RA   Das T., Purkayastha-Mukherjee C., D'Angelo J., Weir M.;
RT   "A conserved F-box gene with unusual transcript localization.";
RL   Dev. Genes Evol. 212:134-140(2002).
CC   -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC       proteins and promotes their ubiquitination and degradation.
CC   -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 F-box domain.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AK084506; BAC39201.1; -.
DR   EMBL; BC006913; AAH06913.1; -.
DR   EMBL; BC021329; AAH21329.1; -.
DR   EMBL; AF467463; AAL75967.1; -.
DR   MGD; MGI:2141676; Fbxl14.
DR   InterPro; IPR001810; F-box.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007089; LRR_cys.
DR   InterPro; IPR008945; Skp1_Skp2.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00560; LRR; 6.
DR   SMART; SM00256; FBOX; 1.
DR   PROSITE; PS50181; FBOX; FALSE_NEG.
KW   Leucine-rich repeat; Repeat; Ubl conjugation pathway.
FT   DOMAIN        2     48       F-box.
FT   REPEAT       91    120       LRR 1.
FT   REPEAT      170    194       LRR 2.
FT   REPEAT      203    231       LRR 3.
FT   REPEAT      254    280       LRR 4.
FT   REPEAT      331    356       LRR 5.
FT   REPEAT      357    381       LRR 6.
FT   CONFLICT     22     22       V -> F (in Ref. 2; AAH21329).
SQ   SEQUENCE   400 AA;  43864 MW;  E0B297E4B4F83C22 CRC64;

  Query Match             86.4%;  Score 38;  DB 1;  Length 400;
  Best Local Similarity   77.8%;  Pred. No. 21;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||||:
```

```
Db        367 TRITKRGLE 375


RESULT 4
FXLE_HUMAN
ID   FXLE_HUMAN     STANDARD;      PRT;   418 AA.
AC   Q8N1E6;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   F-box/LRR-repeat protein 14 (F-box and leucine-rich repeat protein
DE   14).
GN   Name=FBXL14;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC       proteins and promotes their ubiquitination and degradation.
CC   -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 F-box domain.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; BC028132; AAH28132.1; -.
DR   Genew; HGNC:28624; FBXL14.
DR   InterPro; IPR001810; F-box.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007089; LRR_cys.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00560; LRR; 6.
DR   SMART; SM00256; FBOX; 1.
DR   PROSITE; PS50181; FBOX; FALSE_NEG.
KW   Leucine-rich repeat; Repeat; Ubl conjugation pathway.
FT   DOMAIN        2     48       F-box.
FT   REPEAT       91    120       LRR 1.
FT   REPEAT      170    194       LRR 2.
FT   REPEAT      203    231       LRR 3.
FT   REPEAT      254    280       LRR 4.
FT   REPEAT      331    356       LRR 5.
FT   REPEAT      357    381       LRR 6.
SQ   SEQUENCE   418 AA;  45886 MW;  5779961C8177779F CRC64;

  Query Match             86.4%;  Score 38;  DB 1;  Length 418;
  Best Local Similarity   77.8%;  Pred. No. 22;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||||:
Db        367 TRITKRGLE 375


RESULT 5
Q75TW9
ID   Q75TW9      PRELIMINARY;      PRT;   423 AA.
AC   Q75TW9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653.
GN   ORFNames=BH9065301;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH-101;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH-101;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH-101;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126496; BAD18188.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   423 AA;  48264 MW;  72B7EDDE480E9BA0 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 423;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 6
BAD18188
ID   BAD18188    PRELIMINARY;      PRT;   423 AA.
AC   BAD18188;
DT   01-JUN-2004 (TrEMBLrel. 27, Created)
DT   01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653.
GN   BH9065301.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
```

```
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH-101;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH-101;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AH-101;
RA   Takami H., Matsuki A., Takaki Y.;
RT   "Wide-range distribution of insertion sequences identified in B.
RT   halodurans among bacilli and a new transposon disseminated in
RT   alkaliphilic and thermophilic bacilli.";
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126496; BAD18188.1; -.
SQ   SEQUENCE   423 AA;  48264 MW;  72B7EDDE480E9BA0 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 423;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 7
Q75TN6
ID   Q75TN6      PRELIMINARY;      PRT;   427 AA.
AC   Q75TN6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653.
GN   ORFNames=BFI065301;
OS   Bacillus firmus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1399;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM12;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM12;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM12;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126567; BAD18271.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48654 MW;  B812714B694F3C88 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 8
Q75TQ1
ID   Q75TQ1      PRELIMINARY;      PRT;   427 AA.
AC   Q75TQ1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS650 (Transposase of IS653).
GN   ORFNames=BAL065001, BAL065301;
OS   Bacillus alcalophilus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1445;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM485;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM485;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM485;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126554; BAD18256.1; -.
DR   EMBL; AB126557; BAD18259.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48754 MW;  B3B77D309F000033 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 9
Q75TX1
ID   Q75TX1      PRELIMINARY;      PRT;   427 AA.
AC   Q75TX1;
```

```
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653.
GN   ORFNames=BH7065301;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM9774;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM9774;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM9774;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126494; BAD18186.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48724 MW;  332EAB0F1B1815A0 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 10
Q75TX3
ID   Q75TX3      PRELIMINARY;      PRT;   427 AA.
AC   Q75TX3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653.
GN   ORFNames=BH5065301;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM6940;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM6940;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM6940;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126492; BAD18184.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48726 MW;  ED574C872AC089EC CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 11
Q75TX6
ID   Q75TX6      PRELIMINARY;      PRT;   427 AA.
AC   Q75TX6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653.
GN   ORFNames=BH2065301;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A59;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A59;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A59;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126489; BAD18181.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  49016 MW;  D89A03BD5B14AE81 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
```

```
Db        339 TRITKRGRK 347


RESULT 12
Q75TX7
ID   Q75TX7      PRELIMINARY;      PRT;   427 AA.
AC   Q75TX7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS653 (Transposase of IS650).
GN   ORFNames=BH1065001, BH1065301;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM497;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM497;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM497;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126488; BAD18180.1; -.
DR   EMBL; AB126464; BAD18156.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48754 MW;  B3B77D309F000033 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 13
Q75TZ5
ID   Q75TZ5      PRELIMINARY;      PRT;   427 AA.
AC   Q75TZ5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS650.
GN   ORFNames=BH8065001;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126470; BAD18162.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48726 MW;  14692A5DC572A2B8 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 14
Q75TZ7
ID   Q75TZ7      PRELIMINARY;      PRT;   427 AA.
AC   Q75TZ7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS650 (Transposase of IS653).
GN   ORFNames=BH4065001, BH4065301, BH6065001, BH6065301, BH7065001;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM8718, DSM9774, and DSM6939;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM8718, DSM9774, and DSM6939;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM8718, DSM9774, and DSM6939;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126468; BAD18160.1; -.
DR   EMBL; AB126469; BAD18161.1; -.
DR   EMBL; AB126491; BAD18183.1; -.
DR   EMBL; AB126493; BAD18185.1; -.
DR   EMBL; AB126466; BAD18158.1; -.
```

```
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48696 MW;  F14A8D5B2B009CB8 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


RESULT 15
Q75TZ8
ID   Q75TZ8      PRELIMINARY;      PRT;   427 AA.
AC   Q75TZ8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transposase of IS650.
GN   ORFNames=BH5065001;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM6940;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM6940;
RX   PubMed=11418576;
RA   Takami H., Han C., Takaki Y., Ohtsubo E.;
RT   "Identification and distribution of new insertion sequences in the
RT   genome of alkaliphilic Bacillus halodurans C-125.";
RL   J. Bacteriol. 183:4345-4356(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM6940;
RA   Takami H., Matsuki A., Takaki Y.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB126467; BAD18159.1; -.
DR   InterPro; IPR003346; Transposase_20.
DR   InterPro; IPR002525; Transposase_9.
DR   Pfam; PF02371; Transposase_20; 1.
DR   Pfam; PF01548; Transposase_9; 1.
SQ   SEQUENCE   427 AA;  48694 MW;  79A190E26EC35384 CRC64;

  Query Match             81.8%;  Score 36;  DB 2;  Length 427;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TRLTKRGLK 9
              ||:|||| |
Db        339 TRITKRGRK 347


Search completed: December 29, 2004, 12:37:39
Job time : 53.6793 secs
```

THIS PAGE IS BLANK